

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
20	6	5089	4679	gi 511839	ORF1 [Staphylococcus bacteriophage phi 11]	100	100	411
149	3	2032	1577	pir B49703 B497	int gene activator RtnA - bacteriophage phi 11	100	100	456
149	5	2109	1912	gi 166161	Bacteriophage phi-11 int gene activator [Staphylococcus acteriophage phi 11]	100	100	198
349	2	558	409	gi 166159	Integrase (int) [Staphylococcus bacteriophage phi 11]	100	100	150
398	1	1372	707	gi 166159	Integrase (int) [Staphylococcus bacteriophage phi 11]	100	99	666
398	2	783	1001	gi 455128	excisionase (xis) [Staphylococcus bacteriophage phi 11]	100	100	219
502	4	1914	1744	gi 1204912	H. influenzae predicted coding region HI0660 [Haemophilus influenzae]	100	71	171
849	1	2	262	gi 1373002	polyprotein [Bean common mosaic virus]	100	46	261
1349	1	277	140	gi 143359	protein synthesis initiation factor 2 (infB) [Bacillus subtilis] gi 49319	100	82	138
2880	1	21	308	gi 862933	protein kinase C inhibitor-I [Homo sapiens]	100	98	288
3085	1	428	216	gi 1354211	PER112-like protein [Bacillus subtilis]	100	100	213
4168	2	571	398	gi 1354211	PER112-like protein [Bacillus subtilis]	100	100	174
331	1	2	247	gi 426473	nusG gene product [Staphylococcus carnosus]	98	95	246
207	2	1272	1463	gi 460259	enolase [Bacillus subtilis]	97	90	192
331	2	395	850	gi 581638	L11 protein [Staphylococcus carnosus]	97	93	456
366	1	39	215	gi 166161	Bacteriophage phi-11 int gene activator [Staphylococcus acteriophage phi 11]	97	95	177
680	3	718	936	gi 426473	nusG gene product [Staphylococcus carnosus]	97	97	219
3578	1	284	144	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	97	79	141
157	1	321	518	gi 1022726	unknown [Staphylococcus haemolyticus]	96	88	198
205	133	116470	16147	gi 1165302	ISI0 [Bacillus subtilis]	96	91	324
3919	1	48	401	gi 871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	96	81	354
4133	1	830	417	gi 1022726	unknown [Staphylococcus haemolyticus]	96	84	414
4168	1	708	355	gi 1354211	PER112-like protein [Bacillus subtilis]	96	95	354
4207	1	312	157	gi 602031	similar to trimethylamine DH [Mycoplasma capricolum] pir S49950 S49950	96	86	156
					probable trimethylamine dehydrogenase (EC .5.99.7) - Mycoplasma capricolum (SGC3) (fragment)			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4227	2	152	331	gi1871784	Cip-like ATP-dependent protease binding subunit [Bos taurus]	96	81	180
4416	1	570	286	gi11022726	unknown [Staphylococcus haemolyticus]	96	84	285
22	1	858	430	gi1511070	UreG [Staphylococcus xylosus]	95	88	429
22	7	4362	4036	gi1581787	urease gamma subunit [Staphylococcus xylosus]	95	79	327
82	6	8794	9114	pirJG0008 JG00	ribosomal protein S7 - Bacillus stearothermophilus	95	83	321
154	9	9280	7838	gi11354211	PET112-like protein [Bacillus subtilis]	95	92	1443
186	3	2798	2055	gi11514656	serine O-acetyltransferase [Staphylococcus xylosus]	95	87	744
205	5	4406	4014	gi1142462	ribosomal protein S11 [Bacillus subtilis]	95	85	393
205	7	5017	4793	gi1142459	initiation factor 1 [Bacillus subtilis]	95	84	225
205	21	11365	10991	gi11044974	ribosomal protein L14 [Bacillus subtilis]	95	93	375
259	5	7288	6644	sp P47995 YSEA_HYPOTHETICAL PROTEIN IN SECA 5'REGION (ORF1) (FRAGMENT).		95	85	645
302	3	795	1097	gi140186	homologous to E.coli ribosomal protein L27 [Bacillus subtilis] i1143592 L27	95	89	303
					ribosomal protein [Bacillus subtilis] ir C21895 C21895 ribosomal protein L27 - Bacillus subtilis p P05657 RL27_BACSU 50S RIBOSOMAL PROTEIN L27			
					(BL30) (BL24). i140175 L24 gene prod			
310	1	579	1523	gi11177684	chorismate mutase [Staphylococcus xylosus]	95	92	945
414	1	2	163	pir C48396 C483	ribosomal protein L34 - Bacillus stearothermophilus	95	90	162
4185	2	125	277	gi11276841	glutamate synthase (GOGAT) [Porphyra purpurea]	95	86	153
22	2	1028	723	gi1511069	UreF [Staphylococcus xylosus]	94	91	306
22	5	5046	3310	gi1410516	urease alpha subunit [Staphylococcus xylosus]	94	85	1737
60	4	815	1372	gi1666116	glucose kinase [Staphylococcus xylosus]	94	87	558
205	18	10012	9536	gi11044978	ribosomal protein S8 [Bacillus subtilis]	94	78	477
326	4	3378	2542	gi1557492	dihydroxynaphthoic acid (DHNA) synthetase [Bacillus subtilis] gi1143186	94	85	837
					dihydroxynaphthoic acid (DHNA) synthetase [Bacillus ubtilis]			
414	3	737	955	gi1467386	thiophen and furan oxidation [Bacillus subtilis]	94	77	219
426	3	2260	1823	gi11263908	putative [Staphylococcus epidermidis]	94	87	438
534	1	2	355	gi1633650	enzyme II(mannitol) [Staphylococcus carnosus]	94	84	354
1017	1	2	229	gi1149435	putative [Lactococcus lactis]	94	73	228
3098	1	330	184	gi1413952	ipa-28d gene product [Bacillus subtilis]	94	50	147

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3232	1	630	316	gi 1022725	unknown [Staphylococcus haemolyticus]	94	84	315
42	5	2089	2259	pir B48396 B483	ribosomal protein L33 - Bacillus stearothermophilus	93	81	171
101	2	1745	1383	gi 155345	arsenic efflux pump protein [Plasmid pSX267]	93	82	363
205	124	11227	11865	sp P14577 RL16_	50S RIBOSOMAL PROTEIN L16.	93	83	363
259	4	8291	5673	gi 499335	secA protein [Staphylococcus carnosus]	93	85	2619
275	1	2226	1114	gi 633650	enzyme II(mannitol) [Staphylococcus carnosus]	93	86	1113
444	6	6207	5773	gi 1022726	unknown [Staphylococcus haemolyticus]	93	81	435
491	1	152	622	gi 46912	ribosomal protein L13 [Staphylococcus carnosus]	93	88	471
607	6	1674	2033	gi 1022726	unknown [Staphylococcus haemolyticus]	93	83	360
653	1	973	488	gi 580890	translation initiation factor IF3 (AA 1-172) [Bacillus tearothermophilus]	93	77	486
1864	1	3	194	gi 306553	ribosomal protein small subunit [Homo sapiens]	93	93	192
2997	1	28	300	gi 143390	carbamy1 phosphate synthetase [Bacillus subtilis]	93	82	273
3232	2	907	596	gi 1022725	unknown [Staphylococcus haemolyticus]	93	84	312
3761	2	794	621	gi 1022725	unknown [Staphylococcus haemolyticus]	93	88	174
16	1	3	374	gi 142781	putative cytoplasmic protein; putative [Bacillus subtilis]	92	83	372
31	7	5915	6124	gi 1136430	KIAA0185 protein [Homo sapiens]	92	46	210
56	19	126483	127391	gi 467401	unknown [Bacillus subtilis]	92	80	909
69	6	5882	6130	gi 530200	trophoblastin [Ovis aries]	92	53	249
145	3	2568	2038	gi 1022725	unknown [Staphylococcus haemolyticus]	92	80	531
171	3	2760	2362	gi 517475	D-amino acid transaminase [Staphylococcus haemolyticus]	92	86	399
205	112	7495	6962	gi 49189	secY gene product [Staphylococcus carnosus]	92	85	534
205	119	110812	110255	gi 1044976	ribosomal protein L5 [Bacillus subtilis]	92	82	558
219	1	710	357	gi 1303812	YqeV [Bacillus subtilis]	92	88	354
344	3	1575	1805	gi 1405474	CspC protein [Bacillus cereus]	92	85	231
699	1	20	361	gi 413999	ipa-75d gene product [Bacillus subtilis]	92	81	342
1343	1	2	160	pir A45434 A454	ribosomal protein L19 - Bacillus stearothermophilus	92	84	159

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1958	1	524	264	gi 1407908	ElIscr [Staphylococcus xylosus]	92	80	261
3578	2	718	386	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	92	78	333
3385	1	644	324	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	92	81	321
3640	1	4	402	gi 1022726	unknown [Staphylococcus haemolyticus]	92	81	399
4362	1	14	178	gi 1450688	hsdM gene of EcoprrI gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	92	78	165
4446	1	358	182	gi 1022725	unknown [Staphylococcus haemolyticus]	92	82	177
4549	1	462	232	gi 1022726	unknown [Staphylococcus haemolyticus]	92	80	231
4626	1	3	224	gi 1022725	unknown [Staphylococcus haemolyticus]	92	84	222
2	4	3980	4531	gi 1535349	CodW [Bacillus subtilis]	91	74	552
28	1	2	1126	gi 1001376	hypothetical protein [Synecocystis sp.]	91	78	1125
60	5	1354	1701	gi 1226043	orf2 downstream of glucose kinase [Staphylococcus xylosus]	91	80	348
101	1	1989	1036	gi 150728	arsenic efflux pump protein [Plasmid pI258]	91	80	954
187	2	412	1194	gi 142559	ATP synthase alpha subunit [Bacillus megaterium]	91	79	783
205	22	11579	11298	gi 140149	SL7 protein (AA 1-87) [Bacillus subtilis]	91	83	282
206	7	8184	10262	gi 1072418	glcA gene product [Staphylococcus carnosus]	91	83	2079
306	2	3885	2326	gi 143012	GMP synthetase [Bacillus subtilis]	91	78	1560
306	3	5319	3826	gi 1467399	IMP dehydrogenase [Bacillus subtilis]	91	79	1494
310	3	2194	3207	gi 1177685	ccpA gene product [Staphylococcus xylosus]	91	81	1014
343	4	2974	3150	gi 1949974	sucrose repressor [Staphylococcus xylosus]	91	82	177
480	3	1606	3042	gi 1433991	ATP synthase subunit beta [Bacillus subtilis]	91	85	1437
536	3	2026	1280	gi 143366	adenylosuccinate lyase (PUR-B) [Bacillus subtilis] pir C29326 WZBSDS adenylosuccinate lyase (EC 4.3.2.2) - Bacillus ubtilis	91	79	747
532	1	1064	615	gi 297874	fructose-bisphosphate aldolase [Staphylococcus carnosus] pir A49943 A49943 fructose-bisphosphate aldolase (EC 4.1.2.13) - taphylococcus carnosus (strain TM300)	91	79	450
637	1	1	1536	gi 143597	CTP synthetase [Bacillus subtilis]	91	79	1536
859	1	21	359	gi 385178	unknown [Bacillus subtilis]	91	66	339

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1327	1	339	530	gi 496558	orfX [Bacillus subtilis]	91	71	192
2515	1	466	275	gi 511070	UreG [Staphylococcus xylosus]	91	85	192
2594	1	2	202	gi 146824	beta-cystathionase [Escherichia coli]	91	75	201
3764	1	847	425	gi 1022725	unknown [Staphylococcus haemolyticus]	91	78	423
4011	1	127	495	gi 1022726	unknown [Staphylococcus haemolyticus]	91	79	369
4227	1	1	177	gi 296464	ATPase [Lactococcus lactis]	91	66	177
42	3	815	1033	gi 520401	catalase [Haemophilus influenzae]	90	86	219
51	8	3717	4607	gi 580899	OppF gene product [Bacillus subtilis]	90	74	891
129	3	5317	4001	gi 1146206	glutamate dehydrogenase [Bacillus subtilis]	90	76	1317
164	17	116628	16933	spiP05766 RS15_30S RIBOSOMAL PROTEIN S15 (BS18).		90	74	306
171	5	2983	2819	gi 517475	D-amino acid transaminase [Staphylococcus haemolyticus]	90	78	165
205	4	4497	3550	gi 142463	RNA polymerase alpha-core-subunit [Bacillus subtilis]	90	76	948
205	6	4748	4410	gi 1044989	ribosomal protein S13 [Bacillus subtilis]	90	73	339
205	10	7165	6404	gi 49189	secY gene product [Staphylococcus carnosus]	90	81	762
205	11	6645	6472	gi 49189	secY gene product [Staphylococcus carnosus]	90	78	174
205	127	113692	13345	gi 786157	Ribosomal Protein S19 [Bacillus subtilis]	90	79	348
205	131	115858	15496	gi 1165303	L3 [Bacillus subtilis]	90	79	363
260	5	7023	5773	gi 1161380	IcaA [Staphylococcus epidermidis]	90	78	1251
299	6	3378	3947	gi 467440	'phosphoribosylpyrophosphate synthetase [Bacillus subtilis] gi 40218 PRPP synthetase (AA 1-317) [Bacillus subtilis]	90	78	570
320	2	1025	1717	gi 312443	carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus aldolyticus]	90	75	693
330	4	1581	1769	gi 986963	beta-tubulin [Sporidiobolus pararoseus]	90	80	189
369	1	954	523	pir S34762 S347	L-serine dehydratase beta chain - Clostridium sp.	90	77	432
557	1	3	188	gi 1511589	M. jannaschii predicted coding region MJ1624 [Methanococcus jannaschii]	90	54	186
663	2	667	1200	gi 143786	tryptophanyl-tRNA synthetase (EC 6.1.1.2) [Bacillus subtilis] pir JT0481 YWBS tryptophan--tRNA ligase (EC 6.1.1.2) - Bacillus ubtilis	90	73	534
717	1	1	261	gi 143065	hubst [Bacillus stearothermophilus]	90	79	261
745	4	1059	865	gi 1205433	H. influenzae predicted coding region HI1190 [Haemophilus influenzae]	90	81	195

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1007	1	386	565	gi 143366	adenylosuccinate lyase (PUR-B) [Bacillus subtilis] pir C29326 WZBSDS	90	77	180
					adenylosuccinate lyase (EC 4.3.2.2) - Bacillus ubtilis			
1054	1	579	331	gi 1033122	ORF_f729 [Escherichia coli]	90	50	249
1156	1	117	707	gi 1477776	CipP [Bacillus subtilis]	90	80	591
1180	1	408	205	gi 1377831	unknown [Bacillus subtilis]	90	74	204
1253	1	1	462	gi 40046	phosphoglucose isomerase A (AA 1-449) [Bacillus stearothermophilus]	90	75	462
					ir S15936 NUBSSA glucose-6-phosphate isomerase (EC 5.3.1.9) A - cillus			
					stearothermophilus			
2951	1	3	269	gi 144816	formyltetrahydrofolate synthetase (FTHFS) (ttg start codon) (EC .3.4.3)	90	76	267
					[Moorella thermoacetica]			
3140	1	327	166	gi 1070014	protein-dependent [Bacillus subtilis]	90	52	162
4594	1	3	233	gi 871784	Cip-like ATP-dependent protease binding subunit [Bos taurus]	90	76	231
87	1	1028	1750	gi 467327	unknown [Bacillus subtilis]	89	75	723
112	1	2	505	gi 153741	ATP-binding protein [Streptococcus mutans]	89	77	504
118	1	120	398	gi 1303804	YqeQ [Bacillus subtilis]	89	75	279
128	4	3545	3757	gi 460257	triose phosphate isomerase [Bacillus subtilis]	89	84	213
164	12	11667	12755	gi 39954	IF2 (aa 1-741) [Bacillus stearothermophilus]	89	80	1089
205	13	7875	7405	gi 216338	ORF for L15 ribosomal protein [Bacillus subtilis]	89	76	471
205	32	16152	15823	gi 1165303	L3 [Bacillus subtilis]	89	80	330
270	3	2407	2207	pir C41902 C419	arsenate reductase (EC 1.-.-.-) - Staphylococcus xylosus plasmid pSX267	89	81	201
395	2	157	672	gi 520574	glutamate racemase [Staphylococcus haemolyticus]	89	80	516
494	1	3	839	gi 396259	protease [Staphylococcus epidermidis]	89	77	837
510	1	1	444	gi 40046	phosphoglucose isomerase A (AA 1-449) [Bacillus stearothermophilus]	89	74	444
					ir S15936 NUBSSA glucose-6-phosphate isomerase (EC 5.3.1.9) A - cillus			
					stearothermophilus			
615	1	2124	1210	gi 1303812	YqeV [Bacillus subtilis]	89	74	915
841	1	18	341	gi 1165303	L3 [Bacillus subtilis]	89	80	324
1111	1	352	813	gi 47146	thermonuclease [Staphylococcus intermedius]	89	70	462
1875	1	2	256	gi 1205108	ATP-dependent protease binding subunit [Haemophilus influenzae]	89	82	255
2963	1	11	367	gi 467458	cell division protein [Bacillus subtilis]	89	83	357

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3020	1	90	362	gi 1239988	hypothetical protein [Bacillus subtilis]	89	66	273
3565	1	2	400	gi 1256635	dihydroxy-acid dehydratase [Bacillus subtilis]	89	75	399
3586	1	105	314	gi 580832	ATP synthase subunit gamma [Bacillus subtilis]	89	82	210
3629	1	794	399	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	89	78	396
3688	1	2	400	gi 1146206	glutamate dehydrogenase [Bacillus subtilis]	89	75	399
3699	1	794	399	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	89	75	396
4016	1	428	216	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	89	71	213
4177	1	471	301	gi 149426	putative [Lactococcus lactis]	89	76	171
4436	1	601	302	gi 1022725	unknown [Staphylococcus haemolyticus]	89	80	300
4635	1	320	162	gi 1022725	unknown [Staphylococcus haemolyticus]	89	73	159
2	2	1330	2676	gi 520754	putative [Bacillus subtilis]	88	76	1347
42	2	468	848	sp P42321 CATA_	CATALASE (EC 1.11.1.6)	88	76	381
53	5	6389	4722	gi 474177	alpha-D-1,4-glucosidase [Staphylococcus xylosus]	88	80	1668
56	16	18018	18617	gi 467411	recombination protein [Bacillus subtilis]	88	77	600
60	3	376	843	gi 666116	glucose kinase. [Staphylococcus xylosus]	88	77	468
70	2	1583	1245	gi 44095	replication initiator protein [Listeria monocytogenes]	88	74	339
82	8	11514	12719	pir A60663 A606	translation elongation factor Tu - Bacillus subtilis	88	79	1206
103	7	4179	4391	gi 167181	serine/threonine kinase receptor [Brassica napus]	88	77	213
114	8	7732	8232	gi 1022726	unknown [Staphylococcus haemolyticus]	88	72	501
118	2	308	2011	gi 1303804	YqeQ [Bacillus subtilis]	88	77	1704
141	3	657	1136	gi 1405446	transketolase [Bacillus subtilis]	88	72	480
148	7	5871	6116	gi 1118002	dihydropteroate synthase [Staphylococcus haemolyticus]	88	78	246
165	3	1428	2231	gi 40053	phenylalananyl-tRNA synthetase alpha subunit [Bacillus subtilis]	88	80	804
				ir S11730 YFBSA	phenylalanine--tRNA ligase (EC 6.1.1.20) alpha ain - Bacillus subtilis			
205	28	15027	14185	gi 1165306	L2 [Bacillus subtilis]	88	82	843
225	1	1569	898	gi 1303840	Yqfs [Bacillus subtilis]	88	78	672
235	1	2	1975	gi 452309	valyl-tRNA synthetase [Bacillus subtilis]	88	76	1974

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
339	3	2060	1566	gi 1118002	dihydropteroate synthase [Staphylococcus haemolyticus]	88	73	495
443	4	4325	2928	gi 558559	pyrimidine nucleoside phosphorylase [Bacillus subtilis]	88	73	1398
532	1	3	419	gi 143797	valyl-tRNA synthetase [Bacillus stearothermophilus] sp P11931 SYV_BACST VALYL-TRNA SYNTHETASE (EC 6.1.1.9) VALINE--TRNA LIGASE) (VALRS).	88	78	417
534	3	2504	2968	gi 153049	mannitol-specific enzyme-III [Staphylococcus carnosus] pir JQ0088 JQ0088 phosphotransferase system enzyme II (EC 7.1.69), mannitol-specific, factor III - Staphylococcus carnosus sp PI7876 PTMA_STACA PTS SYSTEM, MANNITOL-SPECIFIC IIA COMPONENT EIIA-MTL) (88	82	465
705	2	584	399	gi 710018	nitrite reductase (nirB) [Bacillus subtilis]	88	70	186
1000	2	1824	1309	gi 1022726	unknown [Staphylococcus haemolyticus]	88	78	516
1299	1	587	324	gi 401786	phosphomannomutase [Mycoplasma pirum]	88	55	264
1341	2	170	400	gi 39963	ribosomal protein L20 (AA 1-119) [Bacillus stearothermophilus] ir S05348 R5BS20 ribosomal protein L20 - Bacillus earothermophilus	88	82	231
1386	1	41	214	pir B47154 B471	signal recognition particle 54K chain homolog Ffh - Bacillus subtilis	88	71	174
1386	2	183	533	pir B47154 B471	signal recognition particle 54K chain homolog Ffh - Bacillus subtilis	88	73	351
2949	1	704	399	gi 535350	CodX [Bacillus subtilis]	88	73	306
2984	1	5	169	gi 218277	O-acetylserine(thiol) lyase [Spinacia oleracea]	88	70	165
3035	1	1	138	gi 493083	dihydroxyacetone kinase [Citrobacter freundii]	88	67	138
3089	1	3	152	gi 606055	ORF_f746 [Escherichia coli]	88	88	150
3917	1	817	410	gi 143378	pyruvate decarboxylase (E-1) beta subunit [Bacillus subtilis] gi 1377836 pyruvate decarboxylase E-1 beta subunit [Bacillus ubtilis]	88	77	408
4199	1	680	342	gi 1405454	aconitase [Bacillus subtilis]	88	82	339
4201	1	734	369	gi 515938	glutamate synthase (ferredoxin) [Synechocystis sp.] pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	88	84	366
4274	1	1	336	gi 515938	glutamate synthase (ferredoxin) [Synechocystis sp.] pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	88	84	336
4308	1	794	399	gi 1146206	glutamate dehydrogenase [Bacillus subtilis]	88	71	396
2	5	4570	6000	gi 535350	CodX [Bacillus subtilis]	87	70	1431
52	8	6781	6482	gi 1064791	function umknown [Bacillus subtilis]	87	66	300

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
73	3	1584	2480	gi 1142992	glycerol kinase (glpK) (EC 2.7.1.30) [Bacillus subtilis] pir B45868 B45868	87	72	897
					glycerol kinase (EC 2.7.1.30) - Bacillus subtilis sp P18157 GLPK_BACSU			
					GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL -PHOSPHOTRANSFERASE)			
					(GLYCEROKINASE) (GK).			
98	12	8813	9100	gi 1467433	unknown [Bacillus subtilis]	87	62	288
124	4	4265	2988	gi 1556886	serine hydroxymethyltransferase [Bacillus subtilis] pir S49363 S49363	87	77	1278
					serine hydroxymethyltransferase - Bacillus ubtilis			
124	6	4457	4032	gi 1556883	Unknown [Bacillus subtilis]	87	66	426
148	5	3741	4559	gi 1467460	unknown [Bacillus subtilis]	87	70	819
164	13	12710	13810	gi 139954	IF2 (aa 1-741) [Bacillus stearothermophilus]	87	72	1101
177	2	1104	2126	gi 1467385	unknown [Bacillus subtilis]	87	78	1023
199	1	1982	1158	gi 1143527	iron-sulfur protein [Bacillus subtilis]	87	77	825
199	2	4717	2933	pir A27763 A277	succinate dehydrogenase (EC 1.3.99.1) flavoprotein - Bacillus subtilis	87	80	1785
205	123	111782	11543	gi 1044972	ribosomal protein L29 [Bacillus subtilis]	87	78	240
205	125	13275	12607	gi 1165309	S3 [Bacillus subtilis]	87	75	669
222	1	2033	1107	gi 1177249	rec233 gene product [Bacillus subtilis]	87	70	927
236	3	1635	1333	gi 11146198	ferredoxin [Bacillus subtilis]	87	80	303
246	5	2585	2292	gi 1467373	ribosomal protein S18 [Bacillus subtilis]	87	77	294
260	2	4189	3422	gi 1161382	IcaC [Staphylococcus epidermidis]	87	72	768
320	3	1696	2391	gi 1312443	carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus aidolyticus]	87	80	696
380	4	1165	1383	gi 142570	ATP synthase c subunit [Bacillus firmus]	87	80	219
414	4	900	1073	gi 1467386	thiophen and furan oxidation [Bacillus subtilis]	87	77	174
425	2	1003	794	gi 1046166	pilin repressor [Mycoplasma genitalium]	87	69	210
448	1	1255	722	gi 405134	acetate kinase [Bacillus subtilis]	87	75	534
480	1	1	711	gi 1142559	ATP synthase alpha subunit [Bacillus megaterium]	87	79	711
481	1	2	352	sp Q06797 RL1_B	50S RIBOSOMAL PROTEIN L1 (BL1).	87	72	351
677	2	359	955	gi 1460911	fructose-bisphosphate aldolase [Bacillus subtilis]	87	78	597
677	3	934	1284	gi 1460911	fructose-bisphosphate aldolase [Bacillus subtilis]	87	78	351
876	1	3	452	gi 1146247	asparaginyl-tRNA synthetase [Bacillus subtilis]	87	79	450

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1376	1	426	214	gi 1065555	F46H6.4 gene product [Caenorhabditis elegans]	87	75	213
2206	1	3	374	gi 215098	excisionase [Bacteriophage 154a]	87	72	372
2938	1	3	290	gi 508979	GTP-binding protein [Bacillus subtilis]	87	69	288
3081	2	126	308	gi 467399	IMP dehydrogenase [Bacillus subtilis]	87	72	183
3535	1	3	401	gi 1405454	aconitase [Bacillus subtilis]	87	80	399
4238	1	547	275	gi 603769	HutU protein, urocanase [Bacillus subtilis]	87	73	273
4	8	10427	8736	gi 603769	HutU protein, urocanase [Bacillus subtilis]	86	72	1692
22	6	4190	3738	gi 410515	urase beta subunit [Staphylococcus xylosus]	86	73	453
54	2	2480	1572	gi 289287	UDP-glucose pyrophosphorylase [Bacillus subtilis]	86	70	909
124	3	2336	1713	gi 556887	uracil phosphoribosyltransferase [Bacillus subtilis] pir S49364 S49364 uracil phosphoribosyltransferase - Bacillus ubtilis	86	74	624
148	3	1349	3448	gi 467458	cell division protein [Bacillus subtilis]	86	75	2100
148	4	3638	3859	gi 467460	unknown [Bacillus subtilis]	86	73	222
152	3	1340	2086	gi 1377835	pyruvate decarboxylase E-1 alpha subunit [Bacillus subtilis]	86	75	747
164	18	17347	19467	gi 1184680	polynucleotide phosphorylase [Bacillus subtilis]	86	72	2121
180	2	554	1159	gi 143467	ribosomal protein S4 [Bacillus subtilis]	86	80	606
205	3	2966	2592	gi 142464	ribosomal protein L17 [Bacillus subtilis]	86	77	375
205	26	13364	12990	gi 40107	ribosomal protein L22 [Bacillus stearothermophilus] ir S10612 S10612 ribosomal protein L22 - Bacillus earothermophilus	86	75	375
246	7	3463	3140	gi 467375	ribosomal protein S6 [Bacillus subtilis]	86	70	324
299	3	1196	1540	gi 39656	spoVG gene product [Bacillus megaterium]	86	70	345
299	7	3884	4345	gi 467440	'phosphoribosylpyrophosphate synthetase [Bacillus subtilis] gi 40218 PRPP synthetase (AA 1-317) [Bacillus subtilis]	86	78	462
304	5	2170	2523	gi 666983	putative ATP binding subunit [Bacillus subtilis]	86	65	354
310	2	1487	1678	gi 1177684	chorismate mutase [Staphylococcus xylosus]	86	71	192
337	5	2086	3405	gi 487434	isocitrate dehydrogenase [Bacillus subtilis]	86	78	1320
339	2	1489	1109	gi 1118003	dihydroneopterin aldolase [Staphylococcus haemolyticus]	86	77	381
358	2	2124	3440	gi 1146219	28.2% of identity to the Escherichia coli GTP-binding protein Era; putative [Bacillus subtilis]	86	73	1317

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
404	2	1015	2058	gi 1303817	YqfA [Bacillus subtilis]	86	78	1044
581	2	661	452	gi 140056	phoP gene product [Bacillus subtilis]	86	71	210
642	2	338	1075	gi 1176399	EpiF [Staphylococcus epidermidis]	86	72	738
770	1	622	347	gi 1143328	phoP protein (put.); putative [Bacillus subtilis]	86	69	276
865	1	1777	890	gi 1146247	asparaginyl-tRNA synthetase [Bacillus subtilis]	86	74	888
868	2	963	1133	gi 1002911	transmembrane protein [Saccharomyces cerevisiae]	86	69	171
904	1	1	162	gi 1303912	YqhW [Bacillus subtilis]	86	72	162
989	1	35	433	gi 1303993	YqkL [Bacillus subtilis]	86	76	399
1212	1	296	150	gi 1414014	ipa-90d gene product [Bacillus subtilis]	86	70	147
1323	1	2	148	gi 140041	pyruvate dehydrogenase (lipoamide) [Bacillus stearothermophilus]	86	75	147
				gi S10798 DEBSPF	pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) pha chain			
					- Bacillus stearothermophilus			
3085	2	540	310	gi 1354211	PET112-like protein [Bacillus subtilis]	86	86	231
3847	1	1	228	gi 296464	ATPase [Lactococcus lactis]	86	63	228
4487	1	476	240	gi 1022726	unknown [Staphylococcus haemolyticus]	86	73	237
4583	1	372	187	gi 1022725	unknown [Staphylococcus haemolyticus]	86	79	186
25	5	4287	5039	gi 1502421	3-ketoacyl-acyl carrier protein reductase [Bacillus subtilis]	85	64	753
56	121	130627	129395	gi 1408507	pyrimidine nucleoside transport protein [Bacillus subtilis]	85	69	1233
68	2	332	1192	gi 1467376	unknown [Bacillus subtilis]	85	74	861
73	2	880	1707	gi 1142992	glycerol kinase (glpK) (EC 2.7.1.30) [Bacillus subtilis]	85	72	828
					pir B45868 B45868			
					glycerol kinase (EC 2.7.1.30) - Bacillus subtilis sp P18157 GLPK_BACSU			
					GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL -PHOSPHOTRANSFERASE)			
					(GLYCEROKINASE) (GK).			
106	4	1505	3490	gi 1143766	(thrSv) (EC 6.1.1.3) [Bacillus subtilis]	85	74	1986
128	2	1153	2202	gi 311924	glycerldehyde-3-phosphate dehydrogenase [Clostridium pasteurianum]	85	75	1050
					pir S34254 S34254			
					glyceraldehyde-3-phosphate dehydrogenase (EC .2.1.12) -			
					Clostridium pasteurianum			
129	4	6466	5252	gi 11064807	ORTHONINE AMINOTRANSFERASE [Bacillus subtilis]	85	73	1215
138	6	3475	5673	gi 11072419	glcB gene product [Staphylococcus carnosus]	85	74	2199
189	1	2	169	gi 1467385	unknown [Bacillus subtilis]	85	65	168

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
205	115	8624	8106	gi 1044981	ribosomal protein S5 [Bacillus subtilis]	85	75	519
205	120	110928	10596	pir A02819 R5BS	ribosomal protein L24 - Bacillus stearothermophilus	85	72	333
220	6	6490	6101	gi 48980	secA gene product [Bacillus subtilis]	85	66	390
231	4	4877	3159	gi 1002520	MutS [Bacillus subtilis]	85	70	1719
243	9	8013	8783	gi 414011	ipa-87r gene product [Bacillus subtilis]	85	72	771
249	2	5894	3186	gi 1405454	aconitase [Bacillus subtilis]	85	73	2709
302	1	140	475	gi 40173	homolog of E.coli ribosomal protein L21 [Bacillus subtilis] ir S18439 S18439 Ribosomal protein L21 - Bacillus subtilis piP26908 RL21_BACSU 50S RIBOSOMAL PROTEIN L21 (BL20).	85	72	336
333	1	5445	2968	gi 442360	ClpC adenosine triphosphatase [Bacillus subtilis]	85	69	2478
364	6	6082	8196	gi 871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	85	68	2115
448	2	1992	1339	gi 405134	acetate kinase [Bacillus subtilis]	85	68	654
747	1	1251	853	gi 11373157	orf-X; hypothetical protein; Method: conceptual translation supplied by author [Bacillus subtilis]	85	73	399
886	2	159	467	gi 541768	hemin permease [Yersinia enterocolitica]	85	55	309
1089	1	1208	606	pir B47154 B471	signal recognition particle 54K chain homolog Ffh - Bacillus subtilis	85	71	603
1163	1	816	409	gi 304155	diaminopimelate decarboxylase [Bacillus methanolicus] sp P41023 DCDA_BACMT DIAMINOPIMELATE DECARBOXYLASE (EC 4.1.1.20) DAP DECARBOXYLASE).	85	62	408
1924	1	487	251	gi 215098	excisionase [Bacteriophage l54a]	85	73	237
2932	1	776	390	gi 1041099	Pyruvate Kinase [Bacillus licheniformis]	85	71	387
3030	1	3	275	gi 42370	pyruvate formate-lyase (AA 1-760) [Escherichia coli] ir S01788 S01788 formate C-acetyltransferase (EC 2.3.1.54) - cherichia coli	85	74	273
3111	1	595	299	gi 63568	limb deformity protein [Gallus gallus]	85	85	297
3778	1	630	316	gi 1391840	beta-subunit of HDT [Pseudomonas fragi]	85	67	315
3835	1	1	387	gi 1204472	type I restriction enzyme ECOR124/3 I M protein [Haemophilus influenzae]	85	56	387
4042	1	3	386	gi 18178	formate acetyltransferase [Chlamydomonas reinhardtii] ir S24997 S24997 formate C-acetyltransferase (EC 2.3.1.54) - lamydomonas reinhardtii	85	70	384
4053	1	35	340	gi 1204472	type I restriction enzyme ECOR124/3 I M protein [Haemophilus influenzae]	85	56	306
4108	1	2	181	gi 1072418	glcA gene product [Staphylococcus carnosus]	85	61	180
4300	1	575	330	gi 151932	fructose enzyme II [Rhodobacter capsulatus]	85	59	246

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4392	1	627	355	gi11022725	unknown [Staphylococcus haemolyticus]	85	74	273
4408	1	2	235	gi1871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	85	62	234
4430	1	578	291	gi11009366	Respiratory nitrate reductase [Bacillus subtilis]	85	68	288
4555	1	2	253	gi1450688	hsdM gene of Ecoprr1 gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	85	52	252
4611	1	481	242	gi11256635	dihydroxy-acid dehydratase [Bacillus subtilis]	85	65	240
4	10	10061	10591	gi146982	fosB gene product [Staphylococcus epidermidis]	84	68	531
13	2	1348	1172	gi1142450	ahrC protein [Bacillus subtilis]	84	56	177
16	4	1803	4652	gi11277198	DNA repair protein [Deinococcus radiodurans]	84	67	2850
22	3	1535	1128	gi1511069	Uref [Staphylococcus xylosus]	84	73	408
23	7	5055	5306	gi1603320	Yer082p [Saccharomyces cerevisiae]	84	61	252
53	11	11597	11145	gi11303948	YqiW [Bacillus subtilis]	84	68	453
53	12	14059	12770	gi1142613	branched chain alpha-keto acid dehydrogenase E2 [Bacillus subtilis] gi11303944 BfmBB [Bacillus subtilis]	84	71	1290
70	1	1332	982	gi146647	ORF (repE) [Staphylococcus aureus]	84	68	351
73	4	2512	4311	gi1142993	glycerol-3-phosphate dehydrogenase (glpD) (EC 1.1.99.5) [Bacillus ubtilis]	84	74	1800
98	7	4324	6096	gi1467427	methionyl-tRNA synthetase [Bacillus subtilis]	84	66	1773
100	9	9501	8680	gi11340128	ORF1 [Staphylococcus aureus]	84	78	822
117	3	1934	3208	gi11237019	Srb [Bacillus subtilis]	84	68	1275
148	6	4720	5670	gi1467462	cysteine synthetase A [Bacillus subtilis]	84	69	951
152	4	2064	2456	gi1143377	pyruvate decarboxylase (E-1) alpha subunit [Bacillus subtilis] pir B36718 DEBSPA pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) lpha chain - Bacillus subtilis	84	70	393
169	7	3634	3861	gi11001342	hypothetical protein [Synechocystis sp.]	84	66	228
171	4	2992	2657	gi1517475	D-amino acid transaminase [Staphylococcus haemolyticus]	84	71	336
186	6	6941	6216	gi1467475	unknown [Bacillus subtilis]	84	70	726
205	9	6261	5692	gi1216340	ORF for adenylate kinase [Bacillus subtilis]	84	71	570
224	2	915	1391	gi1288269	beta-fructofuranosidase [Staphylococcus xylosus]	84	70	477

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
	251	1	92	gi 11303790	YqeI [Bacillus subtilis]	84	65	297
	282	3	1526	gi 1143040	glutamate-1-semialdehyde 2,1-aminotransferase [Bacillus subtilis] pir D42728 D42728 glutamate-1-semialdehyde 2,1-aminomutase (EC .4.3.8) - Bacillus subtilis	84	75	1311
	307	5	3138	gi 11070014	protein-dependent [Bacillus subtilis]	84	62	180
	320	4	2343	gi 1143390	carbamyl phosphate synthetase [Bacillus subtilis]	84	70	1887
	372	1	3	gi 11022725	unknown [Staphylococcus haemolyticus]	84	70	294
	413	2	2201	gi 11256146	YbbQ [Bacillus subtilis]	84	65	861
	439	1	3	gi 11046173	osmotically inducible protein [Mycoplasma genitalium]	84	53	390
	461	3	1362	gi 140211	threonine synthase (thrC) (AA 1-352) [Bacillus subtilis] ir A25364 A25364 threonine synthase (EC 4.2.99.2) - Bacillus btillis	84	69	909
	487	1	3	gi 11144531	integrin-like protein alpha Intlp [Candida albicans]	84	46	297
	491	2	624	pir S08564 R3BS	ribosomal protein S9 - Bacillus stearothermophilus	84	69	282
	491	3	836	pir S08564 R3BS	ribosomal protein S9 - Bacillus stearothermophilus	84	77	198
	548	1	3	gi 1431231	uracil permease [Bacillus caldolyticus]	84	74	339
	728	2	2701	gi 912445	DNA polymerase [Bacillus caldotenax]	84	68	954
	769	1	3	gi 11510953	cobalamin biosynthesis protein N [Methanococcus jannaschii]	84	38	255
	954	1	308	gi 11405454	aconitase [Bacillus subtilis]	84	57	153
	957	1	3	gi 1143402	recombination protein (ttg start codon) [Bacillus subtilis] gi 1303923 RecN [Bacillus subtilis]	84	68	393
	975	1	3	gi 1885934	ClpB [Synechococcus sp.]	84	70	450
	1585	1	3	gi 1510140	ligoendopeptidase F [Lactococcus lactis]	84	56	255
	2954	1	3	gi 1603769	HutU protein, urocanase [Bacillus subtilis]	84	73	321
	2996	1	650	gi 18178	formate acetyltransferase [Chlamydomonas reinhardtii] ir S24997 S24997 formate C-acetyltransferase (EC 2.3.1.54) - lamydomonas reinhardtii	84	65	303
	3766	1	737	gi 1517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	84	72	363
	4022	1	2	gi 11146206	glutamate dehydrogenase [Bacillus subtilis]	84	54	168
	4058	1	620	gi 1151932	fructose enzyme II [Rhodobacter capsulatus]	84	71	309
	4108	2	106	gi 11072418	glcA gene product [Staphylococcus carnosus]	84	77	246

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4183	1	3	308	gi 603769	HutU protein, urocanase [Bacillus subtilis]	84	72	306
4726	1	55	234	gi 146208	glutamate synthase large subunit (EC 2.6.1.53) [Escherichia coli]	84	73	180
				pir A29617 A29617	glutamate synthase (NADPH) (EC 1.4.1.13) large hain -			
					Escherichia coli			
22	4	2043	1576	gi 393297	urease accessory protein [Bacillus sp.]	83	64	468
53	13	14722	13745	gi 142612	branched chain alpha-keto acid dehydrogenase E1-beta [Bacillus ubtilis]	83	68	978
57	16	13357	12872	gi 143132	lactate dehydrogenase (AC 1.1.1.27) [Bacillus caldolyticus]	83	66	486
				pir B29704 B29704	L-lactate dehydrogenase (EC 1.1.1.27) - Bacillus			
					aldolyticus			
66	3	3119	2274	gi 1303894	YqhM [Bacillus subtilis]	83	63	846
66	5	6118	4643	gi 1212730	YqhK [Bacillus subtilis]	83	68	1476
70	3	1864	1523	gi 44095	replication initiator protein [Listeria monocytogenes]	83	73	342
90	1	377	1429	gi 155571	alcohol dehydrogenase I (adhA) (EC 1.1.1.1) [Zymomonas mobilis]	83	70	1053
				pir A35260 A35260	alcohol dehydrogenase (EC 1.1.1.1) I - Zymomonas obilis			
95	2	708	2162	gi 506381	phospho-beta-glucosidase [Bacillus subtilis]	83	70	1455
137	1	68	694	gi 467391	initiation protein of replicaton [Bacillus subtilis]	83	77	627
140	4	3209	2742	gi 634107	kdpB [Escherichia coli]	83	65	468
142	3	3468	2989	gi 1212776	lumazine synthase (b-subunit) [Bacillus amyloliquefaciens]	83	69	480
161	12	5749	6696	gi 903307	ORF75 [Bacillus subtilis]	83	64	948
164	9	9880	11070	gi 49316	ORF2 gene product [Bacillus subtilis]	83	66	1191
164	14	14148	14546	gi 580902	ORF6 gene product [Bacillus subtilis]	83	60	399
170	2	3144	2467	gi 520844	orf4 [Bacillus subtilis]	83	64	678
186	2	2029	1370	gi 289284	cysteinyI-tRNA synthetase [Bacillus subtilis]	83	72	660
205	14	7822	7607	gi 216337	ORE for L30 ribosomal protein [Bacillus subtilis]	83	74	216
237	6	3683	4540	gi 1510488	imidazoleglycerol-phosphate synthase (cyclase) [Methanococcus jannaschii]	83	60	858
301	1	985	638	gi 467419	unknown [Bacillus subtilis]	83	65	348
302	4	1421	2743	gi 508979	GTP-binding protein [Bacillus subtilis]	83	68	1323
321	4	3933	3571	gi 39844	fumarase (citG) (aa 1-462) [Bacillus subtilis]	83	68	363
367	1	2	352	gi 1039479	ORFU [Lactococcus lactis]	83	54	351

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
387	1	3	662	gi 1806281	DNA polymerase I [Bacillus stearothermophilus]	83	70	660
527	2	916	1566	gi 396259	protease [Staphylococcus epidermidis]	83	67	651
533	1	355	179	gi 142455	alanine dehydrogenase (EC 1.4.1.1) [Bacillus stearothermophilus] pir B34261 B34261 alanine dehydrogenase (EC 1.4.1.1) - Bacillus tearothermophilus	83	66	177
536	4	1617	1438	gi 143366	adenylosuccinate lyase (PUR-B) [Bacillus subtilis] pir C29326 WZBSDS adenylosuccinate lyase (EC 4.3.2.2) - Bacillus ubtilis	83	67	180
652	1	2	859	gi 520753	DNA topoisomerase I [Bacillus subtilis]	83	72	858
774	2	200	361	gi 1522665	M. jannaschii predicted coding region MJEC128 [Methanococcus jannaschii]	83	58	162
897	1	120	296	gi 1064807	ORTHININE AMINOTRANSFERASE [Bacillus subtilis]	83	76	177
1213	1	3	491	gi 289288	lexA [Bacillus subtilis]	83	67	489
2529	1	296	150	gi 143786	tryptophanyl-tRNA synthetase (EC 6.1.1.2) [Bacillus subtilis] pir JT0481 YWBS tryptophan--tRNA ligase (EC 6.1.1.2) - Bacillus ubtilis	83	69	147
2973	1	649	326	gi 1109687	Proz [Bacillus subtilis]	83	58	324
3009	1	728	366	gi 1882532	ORF_o294 [Escherichia coli]	83	65	363
3035	2	45	305	gi 950062	hypothetical yeast protein 1 [Mycoplasma capricolum] pir S48578 S48578 hypothetical protein - Mycoplasma capricolum SGC3) (fragment)	83	59	261
3906	1	67	309	gi 1353197	thioredoxin reductase [Eubacterium acidaminophilum]	83	61	243
4458	1	540	271	gi 397526	clumping factor [Staphylococcus aureus]	83	78	270
4570	1	444	223	gi 1022726	unknown [Staphylococcus haemolyticus]	83	74	222
4654	1	97	261	gi 1072419	glcB gene product [Staphylococcus carnosus]	83	79	165
16	2	295	1191	gi 153854	uvs402 protein [Streptococcus pneumoniae]	82	67	897
16	3	1193	1798	gi 153854	uvs402 protein [Streptococcus pneumoniae]	82	70	606
38	12	9644	8724	gi 1204400	N-acetylneuraminate lyase [Haemophilus influenzae]	82	58	921
42	4	988	2019	gi 1841192	catalase [Bacteroides fragilis]	82	70	1032
51	6	2590	3489	gi 143607	sporulation protein [Bacillus subtilis]	82	69	900
56	11	12270	13925	gi 39431	oligo-1,6-glucosidase [Bacillus cereus]	82	60	1656
56	15	17673	18014	gi 467410	unknown [Bacillus subtilis]	82	66	342
61	2	881	3313	gi 143148	transfer RNA-Leu synthetase [Bacillus subtilis]	82	70	2433

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
82	7	9162	11318	gi 48240	elongation factor G (AA 1-691) [Thermus aquaticus thermophilus] ir S15928 EFTWG translation elongation factor G - Thermus aquaticus pip13551 EFG_THETH ELONGATION FACTOR G (EF-G).	82	64	2157
85	2	5470	3260	gi 143369	phosphoribosylformyl glycinamide synthetase II (PUR-Q) [Bacillus ubtilis]	82	66	2211
102	6	3662	5380	gi 1256635	dihydroxy-acid dehydratase [Bacillus subtilis]	82	65	1719
117	4	3242	3493	pir A47154 A471	orf1 5' of Ffh - Bacillus subtilis	82	53	252
128	6	4377	5933	gi 460258	phosphoglycerate mutase [Bacillus subtilis]	82	66	1557
129	2	1229	2182	gi 403373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis] pir S37251 S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis	82	62	954
170	1	2	1441	gi 1377831	unknown [Bacillus subtilis]	82	67	1440
177	1	3	1094	gi 467386	thiophen and furan oxidation [Bacillus subtilis]	82	65	1092
184	4	3572	4039	gi 153566	ORF (19K protein) [Enterococcus faecalis]	82	59	468
189	8	4455	4225	gi 1001878	CspL protein [Listeria monocytogenes]	82	73	231
206	19	21366	20707	gi 473916	lipopeptide antibiotics iturin A [Bacillus subtilis] sp P39144 LP14_BACSU LIPOPEPTIDE ANTIBIOTICS ITURIN A AND SURFACTIN IOSYNTHESIS PROTEIN.	82	50	660
221	2	805	1722	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	82	63	918
223	4	3866	3651	gi 439619	[Salmonella typhimurium Is200 insertion sequence from SAR17, artial.), gene product [Salmonella typhimurium]	82	69	216
260	3	5207	4296	gi 1161381	IcaB [Staphylococcus epidermidis]	82	61	912
315	3	4864	2855	gi 143397	quinol oxidase [Bacillus subtilis]	82	67	2010
321	10	8520	7945	gi 142981	ORF5; This ORF includes a region (aa23-103) containing a potential ron- sulphur centre homologous to a region of Rhodospirillum rubrum nd Chromatium vinosum; putative [Bacillus stearothermophilus] pir PQ0299 PQ0299 hypothetical protein 5 (glda 3' region) -	82	62	576
331	3	1055	1342	gi 436574	ribosomal protein L1 [Bacillus subtilis]	82	71	288
370	2	262	618	gi 1303793	YqeL [Bacillus subtilis]	82	59	357
404	4	3053	4024	gi 1303821	YqfE [Bacillus subtilis]	82	68	972
405	4	4440	3073	gi 1303913	YqhX [Bacillus subtilis]	82	67	1368
436	3	4096	2864	gi 149521	tryptophan synthase beta subunit [Lactococcus lactis] pir S35129 S35129 tryptophan synthase (EC 4.2.1.20) beta chain - actococcus lactis subsp. lactis	82	67	1233

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
441	4	3394	2573	gi 142952	glyceraldehyde-3-phosphate dehydrogenase [Bacillus tearothermophilus]	82	67	822
444	12	110415	11227	gi 1204354	spore germination and vegetative growth protein [Haemophilus influenzae]	82	67	813
446	1	3	191	gi 143387	aspartate transcarbamylase [Bacillus subtilis]	82	66	189
462	3	1007	1210	gi 142521	deoxyribodipyrimidine photolyase [Bacillus subtilis] pir A37192 A37192 uvrB protein - Bacillus subtilis sp P14951 UVRC_BACSU EXCINUCLEASE ABC SUBUNIT C.	82	64	204
537	1	1560	784	gi 853767	UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus ubtilis]	82	61	777
680	2	407	700	gi 426472	secE gene product [Staphylococcus carnosus]	82	69	294
724	2	565	386	gi 143373	phosphoribosyl aminimidazole carboxy formyl ornyltransferase/inosine monophosphate cyclohydrolase (PUR-H(J)) Bacillus subtilis]	82	68	180
763	1	422	213	gi 467458	cell division protein [Bacillus subtilis]	82	35	210
818	1	564	283	gi 1064787	function unknown [Bacillus subtilis]	82	69	282
858	1	175	1176	gi 143043	uroporphyrinogen decarboxylase [Bacillus subtilis] pir B47045 B47045 uroporphyrinogen decarboxylase (EC 4.1.1.37) - acillus subtilis	82	71	1002
895	1	3	599	gi 1027507	ATP binding protein [Borrelia burgdorferi]	82	72	597
939	1	10	399	gi 143795	transfer RNA-Tyr synthetase [Bacillus subtilis]	82	60	390
961	1	1	306	gi 577647	gamma-hemolysin [Staphylococcus aureus]	82	69	306
1192	1	307	155	gi 146974	NH3-dependent NAD synthetase [Escherichia coli]	82	71	153
1317	1	49	375	gi 407908	EIIscr [Staphylococcus xylosus]	82	72	327
1341	1	1	150	gi 39962	ribosomal protein L35 (AA 1-66) [Bacillus stearothermophilus]	82	68	150
2990	2	567	349	gi 534855	ATPase subunit epsilon [Bacillus stearothermophilus] sp P42009 ATPE_BACST ATP SYNTHASE EPSILON CHAIN (EC 3.6.1.34).	82	47	219
3024	1	45	224	gi 467402	unknown [Bacillus subtilis]	82	64	180
3045	1	276	139	gi 467335	ribosomal protein L9 [Bacillus subtilis]	82	60	138
3045	2	558	400	gi 467335	ribosomal protein L9 [Bacillus subtilis]	82	82	159
3091	1	474	238	gi 499335	secA protein [Staphylococcus carnosus]	82	78	237
3107	1	416	210	gi 546918	orfY 3' of comK [Bacillus subtilis, E26, Peptide Partial, 140 aa] pir S43612 S43612 hypothetical protein Y - Bacillus subtilis sp P40398 YHXD_BACSU HYPOTHETICAL PROTEIN IN COMK 3'REGION (ORFY FRAGMENT).	82	64	207

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4332	1	2	319	gi 42086	nitrate reductase alpha subunit [Escherichia coli] p P09152 NARG_ECOLI RESPIRATORY NITRATE REDUCTASE 1 ALPHA CHAIN (EC 7.99.4). {SUB 2-1247}	82	75	318
23	3	3275	2574	gi 1199573	sp Sphingomonas sp.	81	64	702
42	1	638	321	gi 466778	lysine specific permease [Escherichia coli]	81	59	318
48	5	4051	4350	gi 1045937	M. genitalium predicted coding region MG246 [Mycoplasma genitalium]	81	62	300
51	4	1578	2579	pir S16649 S166	dc AC protein - Bacillus subtilis	81	55	1002
53	2	364	1494	gi 1303961	YqjJ [Bacillus subtilis]	81	67	1131
53	8	9419	7971	gi 146930	6-phosphogluconate dehydrogenase [Escherichia coli]	81	66	1449
54	9	10757	10119	gi 143016	permease [Bacillus subtilis]	81	65	639
54	10	13360	11786	gi 143015	gluconate kinase [Bacillus subtilis]	81	64	1575
57	17	13983	13366	pir A25805 A258	L-lactate dehydrogenase (EC 1.1.1.27) - Bacillus subtilis	81	74	618
81	2	2708	2217	gi 1222302	NiFU-related protein [Haemophilus influenzae]	81	54	492
86	1	745	374	gi 414017	ipa-93d gene product [Bacillus subtilis]	81	70	372
103	6	6438	4861	gi 971342	nitrate reductase beta subunit [Bacillus subtilis] sp P42176 NARH_BACSU NITRATE REDUCTASE BETA CHAIN (EC 1.7.99.4).	81	64	1578
120	15	10845	12338	gi 1524392	GbsA [Bacillus subtilis]	81	67	1494
128	5	3676	4413	gi 143319	triose phosphate isomerase [Bacillus megaterium]	81	64	738
131	9	10308	9280	gi 299163	alanine dehydrogenase [Bacillus subtilis]	81	68	1029
143	6	6088	5471	gi 439619	[Salmonella typhimurium IS200 insertion sequence from SARA17, artial.], gene product [Salmonella typhimurium]	81	61	618
169	1	43	825	gi 897795	30S ribosomal protein [Pediococcus acidilactici] sp P49668 RS2_PEDAC 30S RIBOSOMAL PROTEIN S2.	81	65	783
230	1	450	226	gi 1125826	short region of weak similarity to tyrosine-protein kinase receptors in a fibronectin type III-like domain [Caenorhabditis elegans]	81	54	225
233	5	2000	2677	gi 467404	unknown [Bacillus subtilis]	81	63	678
241	2	3081	2149	gi 16510	succinate--CoA ligase (GDP-forming) [Arabidopsis thaliana] ir S30579 S30579 succinate--CoA ligase (GDP-forming) (EC 6.2.1.4) pha chain - Arabidopsis thaliana (fragment)	81	69	933
256	1	1	981	pir S09411 S094	spoIIIE protein - Bacillus subtilis	81	65	981
259	3	3752	2691	sp P28367 RF2_B	PROBABLE PEPTIDE CHAIN RELEASE FACTOR 2 (RF-2) (FRAGMENT).	81	65	1062

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
275	2	1728	3581	gi 726480	L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus ubtilis]	81	68	1854
285	1	1466	735	gi 1204844	H. influenzae predicted coding region HI0594 [Haemophilus influenzae]	81	63	732
296	1	99	1406	gi 467328	adenylosuccinate synthetase [Bacillus subtilis]	81	67	1308
302	9	5590	5889	gi 147485	queA [Escherichia coli]	81	64	300
317	2	1137	1376	gi 154961	resolvase [Transposon Tn917]	81	51	240
343	2	1034	1342	gi 405955	yeaD [Escherichia coli]	81	60	309
360	2	1404	2471	gi 1204570	aspartyl-tRNA synthetase [Haemophilus influenzae]	81	67	1068
364	5	6251	5706	gi 1204652	methylated-DNA--protein-cysteine methyltransferase [Haemophilus influenzae]	81	63	546
372	2	1707	1135	gi 467416	unknown [Bacillus subtilis]	81	65	573
392	1	43	603	pir S09411 S094	spoIIIE protein - Bacillus subtilis	81	65	561
404	9	5252	6154	gi 606745	Bex [Bacillus subtilis]	81	65	903
426	2	1727	1119	gi 39453	Manganese superoxide dismutase [Bacillus caldotenax] ir S22053 S22053 superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus ldotenax	81	66	609
480	7	5653	5889	pir C37083 C370	hypothetical protein II (ompH 3' region) - Salmonella typhimurium (fragment)	81	57	237
625	3	1105	2070	gi 1262360	protein kinase PknB [Mycobacterium leprae]	81	56	966
754	2	504	1064	gi 1303902	YqhU [Bacillus subtilis]	81	71	561
842	1	86	430	gi 1405446	transketolase [Bacillus subtilis]	81	68	345
953	1	798	400	gi 1205429	dipeptide transport ATP-binding protein [Haemophilus influenzae]	81	57	399
961	2	252	401	gi 487686	synergohymenotropic toxin [Staphylococcus intermedius] pir S44944 S44944 synergohymenotropic toxin - Staphylococcus ntermedius	81	72	150
1035	1	1	189	gi 1046138	M. genitalium predicted coding region MG423 [Mycoplasma genitalium]	81	43	189
1280	1	670	449	gi 559164	helicase [Autographa californica nuclear polyhedrosis virus] sp P24307 V143_NPVAC HELICASE.	81	43	222
3371	1	68	241	gi 1322245	mevalonate pyrophosphate decarboxylase [Rattus norvegicus]	81	62	174
3715	1	475	239	gi 537137	ORE_f388 [Escherichia coli]	81	58	237
3908	1	2	325	gi 439619	[Salmonella typhimurium IS200 insertion sequence from SARAI7, artial.], gene product [Salmonella typhimurium]	81	68	324
3940	1	3	401	gi 296464	ATPase [Lactococcus lactis]	81	69	399

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3954	1	1	318	gi 1224069	amidase [Moraxella catarrhalis]	81	68	318
4049	1	337	170	gi 603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis]	81	68	168
				gi 603768	HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis]			
4209	1	1	324	gi 403373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis]	81	58	324
				pir S37251	S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis			
4371	1	627	322	gi 216677	indolepyruvate decarboxylase [Enterobacter cloacae] pir S16013	81	72	306
					indolepyruvate decarboxylase (EC 4.1.1.-) - nterobacter cloacae			
4387	1	19	228	gi 460689	TVG [Thermoactinomyces vulgaris]	81	59	210
4391	1	581	306	gi 1524193	unknown [Mycobacterium tuberculosis]	81	67	276
4425	1	3	341	gi 143015	gluconate kinase [Bacillus subtilis]	81	66	339
9	1	1593	847	gi 1064786	function unknown [Bacillus subtilis]	80	62	747
17	1	544	311	gi 559164	helicase [Autographa californica nuclear polyhedrosis virus]	80	40	234
				sp P24307	V143_NPVAC HELICASE.			
45	2	1159	2448	gi 1109684	ProV [Bacillus subtilis]	80	63	1290
45	5	4032	4733	gi 1109687	ProZ [Bacillus subtilis]	80	55	702
54	8	10266	9502	gi 563952	gluconate permease [Bacillus licheniformis]	80	62	765
62	12	8852	7545	gi 854655	Na/H antiporter system [Bacillus alcalophilus]	80	62	1308
62	14	8087	8683	gi 559713	ORF [Homo sapiens]	80	68	597
67	16	113781	14122	gi 305002	ORE_f356 [Escherichia coli]	80	65	342
70	13	11495	10296	gi 1303995	YqkN [Bacillus subtilis]	80	64	1200
98	9	6336	7130	gi 467428	unknown [Bacillus subtilis]	80	68	795
98	10	7294	7833	gi 467430	unknown [Bacillus subtilis]	80	64	540
98	11	7820	8737	gi 467431	high level kasgamycin resistance [Bacillus subtilis]	80	61	918
109	16	14154	14813	gi 580875	ipa-57d gene product [Bacillus subtilis]	80	63	660
112	15	14294	16636	gi 1072361	pyruvate-formate-lyase [Clostridium pasteurianum]	80	65	2343
139	1	1448	726	gi 506699	CapC [Staphylococcus aureus]	80	58	723
139	2	2179	1448	gi 506698	CapB [Staphylococcus aureus]	80	59	732
174	4	3271	2870	gi 1146242	aspartate 1-decarboxylase [Bacillus subtilis]	80	61	402

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
177	3	2102	2842	gi 467385	unknown [Bacillus subtilis]	80	70	741
184	6	6124	5912	gi 161953	85-kDa surface antigen [Trypanosoma cruzi]	80	46	213
186	4	5368	3875	gi 289282	glutamyl-tRNA synthetase [Bacillus subtilis]	80	65	1494
205	130	115796	15140	gi 40103	ribosomal protein L4 [Bacillus stearothermophilus]	80	66	657
207	1	140	1315	gi 460259	enolase [Bacillus subtilis]	80	67	1176
211	3	1078	1590	gi 410131	ORFX7 [Bacillus subtilis]	80	61	513
235	2	1962	2255	gi 143797	valyl-tRNA synthetase [Bacillus stearothermophilus] sp P11931 SYV_BACST VALYL-TRNA SYNTHETASE (EC 6.1.1.9) VALINE--TRNA LIGASE) (VALRS).	80	55	294
239	1	1	1263	gi 143000	proton glutamate symport protein [Bacillus stearothermophilus] pir S26247 S26247 glutamate/aspartate transport protein - Bacillus tearothermophilus	80	59	1263
272	5	2724	2461	gi 709993	hypothetical protein [Bacillus subtilis]	80	54	264
301	3	1446	1111	gi 467418	unknown [Bacillus subtilis]	80	58	336
310	4	5697	4501	gi 1177686	acuC gene product [Staphylococcus xylosus]	80	67	1197
310	6	5258	7006	gi 348053	acetyl-CoA synthetase [Bacillus subtilis]	80	67	1749
310	7	7410	9113	gi 1103865	formyl-tetrahydrofolate synthetase [Streptococcus mutans]	80	67	1704
325	3	1114	1389	gi 310325	outer capsid protein [Rotavirus sp.]	80	40	276
337	1	1268	636	gi 537049	ORF_o470 [Escherichia coli]	80	55	633
374	2	929	1228	gi 1405448	YneF [Bacillus subtilis]	80	70	300
375	5	3062	3331	gi 467448	unknown [Bacillus subtilis]	80	68	270
388	1	267	587	gi 1064791	function unkown [Bacillus subtilis]	80	65	321
394	1	9	659	gi 304976	matches PS00017: ATP GTP-A and PS00301: EFACITOR GTP; similar to longation factor G, TetM/TetO tetracycline-resistance proteins Escherichia coli]	80	65	651
456	1	625	1263	gi 1146183	putative [Bacillus subtilis]	80	65	639
475	1	1	654	gi 288269	beta-fructofuranosidase [Staphylococcus xylosus]	80	66	654
544	2	1449	2240	gi 529754	spec [Streptococcus pyogenes]	80	50	792
622	4	1623	1871	gi 1483545	unknown [Mycobacterium tuberculosis]	80	65	249
719	1	1	1257	gi 1064791	function unkown [Bacillus subtilis]	80	68	1257
739	1	107	838	gi 666983	putative ATP binding subunit [Bacillus subtilis]	80	61	732

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
745	2	581	414	gi11511600	coenzyme PQQ synthesis protein III [Methanococcus jannaschii]	80	61	168
822	1	17	679	gi1410141	ORFX17 [Bacillus subtilis]	80	68	663
827	2	991	836	gi11205301	leukotoxin secretion ATP-binding protein [Haemophilus influenzae]	80	54	156
1044	1	3	149	gi160632	vp2 [Marburg virus]	80	55	147
1220	2	571	413	pirA61072	EPGSG gallidermin precursor - Staphylococcus gallinarum	80	74	159
2519	1	75	275	gi1147556	dpj [Escherichia coli]	80	45	201
2947	1	503	279	gi11184680	polynucleotide phosphorylase [Bacillus subtilis]	80	62	225
3120	1	2	226	gi1517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	80	65	225
3191	1	294	148	gi1151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pirA44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	80	59	147
3560	2	285	434	gi1217130	photosystem I core protein B [Synechococcus vulcanus]	80	70	150
3655	1	47	346	gi1415855	deoxyribose aldolase [Mycoplasma hominis]	80	56	300
3658	2	324	584	gi1551531	2-nitropropane dioxygenase [Williopsis saturnus]	80	54	261
3769	1	798	400	gi11339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	80	68	399
3781	1	692	348	gi1166412	NADH-glutamate synthase [Medicago sativa]	80	62	345
3988	1	48	287	gi11204696	fructose-permease IIBC component [Haemophilus influenzae]	80	69	240
4030	1	571	287	gi11009366	Respiratory nitrate reductase [Bacillus subtilis]	80	60	285
4092	1	547	275	gi11370207	orf6 [Lactobacillus sake]	80	69	273
4103	1	680	342	gi139956	IIIGlc [Bacillus subtilis]	80	65	339
4231	1	692	348	gi1289287	UDP-glucose pyrophosphorylase [Bacillus subtilis]	80	65	345
4265	1	595	299	gi1603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis] gi1603768 HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis	80	63	297
4504	1	498	250	gi11339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	80	68	249
2	6	5998	6798	gi1535351	CodY [Bacillus subtilis]	79	63	801
4	7	8295	7051	gi1603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis] gi1603768 HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis	79	64	1245
25	6	5273	5515	pirA36728 A367	acyl carrier protein - Rhizobium meliloti	79	65	243

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
59	2	1173	1424	gi 147923	threonine dehydratase 2 (EC 4.2.1.16) [Escherichia coli]	79	75	252
60	1	1	204	gi 666115	orf1 upstream of glucose kinase [Staphylococcus xylosus] pir S52351 S52351 hypothetical protein 1 - Staphylococcus xylosus	79	60	204
81	1	3002	1590	gi 466882	ppsl; B1496_C2_189 [Mycobacterium leprae]	79	64	1413
85	7	7023	6505	gi 143364	phosphoribosyl aminimidazole carboxylase I (PUR-E) [Bacillus ubtilis]	79	60	519
89	6	5660	4554	gi 144906	product homologous to E.coli thioredoxin reductase: J.Biol.Chem. 1988) 263:9015-9019, and to F52a protein of alkyl hydroperoxide eductase from S.typhimurium: J.Biol.Chem. (1990) 265:10535-10540; pen reading frame A [Clostridium pasteurianum]	79	35	1107
102	11	7489	8571	gi 143093	ketol-acid reductoisomerase [Bacillus subtilis] sp P37253 ILVC_BACSU KETOL- ACID REDUCTOISOMERASE (EC 1.1.1.86) ACETOHYDROXY-ACID ISOMEROREDUCTASE) (ALPHA-KETO-BETA-HYDROXYLACIL EDUCTOISOMERASE) .	79	64	1083
102	14	11190	12563	gi 149428	putative [Lactococcus lactis]	79	65	1374
127	9	7792	9372	gi 458688	PrfC/RF3 [Dichelobacter nodosus]	79	68	1581
139	3	2540	1983	gi 506697	CapA [Staphylococcus aureus]	79	55	558
144	2	1644	1156	gi 1498296	peptide methionine sulfoxide reductase [Streptococcus pneumoniae]	79	47	489
148	2	529	1098	gi 467457	hypoxanthine-guanine phosphoribosyltransferase [Bacillus subtilis] gi 467457 hypoxanthine-guanine phosphoribosyltransferase [Bacillus ubtilis]	79	59	570
150	1	965	591	gi 755602	unknown [Bacillus subtilis]	79	61	375
176	1	1039	587	gi 297874	fructose-bisphosphate aldolase [Staphylococcus carnosus] pir A49943 A49943 fructose-bisphosphate aldolase (EC 4.1.2.13) - taphylococcus carnosus (strain TM300)	79	65	453
186	7	7584	6874	gi 1314298	ORF5; putative Sms protein; similar to Sms proteins from Haemophilus influenzae and Escherichia coli [Listeria monocytogenes]	79	64	711
205	16	8887	8498	gi 1044980	ribosomal protein L18 [Bacillus subtilis]	79	70	390
211	1	1	519	gi 1303994	YqkM [Bacillus subtilis]	79	62	519
223	2	4183	2801	gi 488430	alcohol dehydrogenase 2 [Entamoeba histolytica]	79	60	1383
243	8	8915	7896	gi 580883	ipa-88d gene product [Bacillus subtilis]	79	60	1020
279	4	3721	4329	gi 413930	ipa-6d gene product [Bacillus subtilis]	79	59	609
300	1	11	1393	gi 403372	glycerol 3-phosphate permease [Bacillus subtilis]	79	62	1383
307	3	2930	1935	gi 950062	hypothetical yeast protein 1 [Mycoplasma capricolum] pir S48578 S48578 hypothetical protein - Mycoplasma capricolum SGC3) (fragment)	79	60	996

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
352	6	110106	8886	gi 216854	P47K [Pseudomonas chlororaphis]	79	59	1221
412	1	1153	578	gi 143177	putative [Bacillus subtilis]	79	51	576
481	3	621	1124	gi 786163	Ribosomal Protein L10 [Bacillus subtilis]	79	66	504
516	1	702	352	gi 805090	NisF [Lactococcus lactis]	79	48	351
525	2	2457	1426	gi 143371	phosphoribosyl aminimidazole synthetase (PUR-M) [Bacillus subtilis] pir H29326 AJBSCL phosphoribosylformylglycinamide cyclo-ligase EC 6.3.3.1) - Bacillus subtilis	79	61	1032
538	4	3448	2825	gi 1370207	orf6 [Lactobacillus sake]	79	67	624
570	1	2	421	gi 476160	arginine permease substrate-binding subunit [Listeria monocytogenes]	79	61	420
645	8	2663	3241	gi 153898	transport protein [Salmonella typhimurium]	79	62	579
683	1	75	374	gi 1064795	function unknown [Bacillus subtilis]	79	62	300
816	3	4700	3987	gi 1407784	orf-1; novel antigen [Staphylococcus aureus]	79	62	714
2929	1	3	401	gi 1524397	glycine betaine transporter OpuD [Bacillus subtilis]	79	61	399
2937	1	357	202	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	79	58	156
2940	1	768	385	gi 149429	putative [Lactococcus lactis]	79	72	384
2946	1	570	286	gi 143267	2-oxoglutarate dehydrogenase (odhA; EC 1.2.4.2) [Bacillus subtilis]	79	61	285
2999	1	3	212	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	79	59	210
3022	1	514	332	gi 450686	3-phosphoglycerate kinase [Thermotoga maritima]	79	61	183
3064	1	3	314	gi 1204436	pyruvate formate-lyase [Haemophilus influenzae]	79	60	312
3083	1	2	220	gi 1149662	hypD gene product [Clostridium perfringens]	79	56	219
3126	1	701	411	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	79	55	291
3181	1	607	326	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	79	59	282
3345	1	3	476	gi 871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	79	63	474
3718	1	536	270	pir C36889 C368	leuB protein, inactive - Lactococcus lactis subsp. lactis (strain IL1403)	79	71	267
3724	2	159	401	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	79	64	243
3836	1	608	312	gi 1524193	unknown [Mycobacterium tuberculosis]	79	65	297
3941	1	2	334	gi 415855	deoxyribose aldolase [Mycoplasma hominis]	79	54	333
4113	1	3	341	gi 143015	gluconate kinase [Bacillus subtilis]	79	63	339

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4501	1	406	209	gi11022726	unknown [Staphylococcus haemolyticus]	79	66	198
4612	1	2	238	gi1460689	TVG [Thermoactinomyces vulgaris]	79	58	237
2	1	2	1213	gi1520753	DNA topoisomerase I [Bacillus subtilis]	78	64	1212
8	2	2266	1220	gi1216151	DNA polymerase (gene L; ttg start codon) [Bacteriophage SPO2] gi1579197 SPO2 DNA polymerase (aa 1-648) [Bacteriophage SPO2] pirA21498 DJBPS2 DNA-directed DNA polymerase (EC 2.7.7.7) - phage PO2	78	72	1047
9	2	1340	1089	gi11064787	function unknown [Bacillus subtilis]	78	57	252
32	8	6803	7702	gi1146974	NH3-dependent NAD synthetase [Escherichia coli]	78	63	900
36	4	2941	3138	gi1290503	glutamate permease [Escherichia coli]	78	53	198
53	15	17684	16221	gi11303941	YqiV [Bacillus subtilis]	78	58	1464
57	14	110520	12067	gi11072418	glcA gene product [Staphylococcus carnosus]	78	65	1548
66	7	6798	5812	gi11212729	YqhJ [Bacillus subtilis]	78	67	987
67	4	4029	4376	gi1466612	nika [Escherichia coli]	78	71	348
91	9	110058	10942	gi1467380	stage 0 sporulation [Bacillus subtilis]	78	50	885
102	12	8574	10130	gi1149426	putative [Lactococcus lactis]	78	61	1557
112	6	3540	4463	gi1854234	cymG gene product [Klebsiella oxytoca]	78	56	924
124	2	1888	1061	gi1405622	unknown [Bacillus subtilis]	78	60	828
130	3	1805	2260	gi11256636	putative [Bacillus subtilis]	78	71	456
133	1	751	377	gi1168060	lamB [Emericella nidulans]	78	59	375
166	4	7125	6163	gi1451216	Mannosephosphate Isomerase [Streptococcus mutans]	78	63	963
186	1	1586	795	gi1289284	cysteinyI-tRNA synthetase [Bacillus subtilis]	78	63	792
195	4	2749	2315	gi11353874	unknown [Rhodobacter capsulatus]	78	58	435
199	3	4279	3623	gi1143525	succinate dehydrogenase cytochrome b-558 subunit [Bacillus subtilis] pirA29843 DEBSSC succinate dehydrogenase (EC 1.3.99.1) cytochrome 558 - Bacillus subtilis	78	57	657
199	4	7209	5557	gi1142521	deoxyribodipyrimidine photolyase [Bacillus subtilis] pirA37192 A37192 uvrB protein - Bacillus subtilis sp P14951 UVRC_BACSU EXCINUCLEASE ABC SUBUNIT C.	78	62	1653
223	3	3831	3523	gi1439596	{Escherichia coli IS200 insertion sequence from ECOR63, partial.}, ene product [Escherichia coli]	78	47	309

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
299	4	1865	2149	gi 467439	temperature sensitive cell division [Bacillus subtilis]	78	62	285
321	9	7734	7315	gi 142979	ORF3 is homologous to an ORF downstream of the spot gene of E.coli; RF3 [Bacillus stearothermophilus]	78	55	420
352	4	3714	3944	gi 349050	actin 1 [Pneumocystis carinii]	78	42	231
352	5	7592	6093	gi 903587	NADH dehydrogenase subunit 5 [Bacillus subtilis] sp P39755 NDHF_BACSU NADH DEHYDROGENASE SUBUNIT 5 (EC 1.6.5.3) NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5).	78	58	1500
376	1	2	583	gi 551693	dethiobiotin synthase [Bacillus sphaericus]	78	34	582
424	2	1595	1768	gi 152417	alpha-acetolactate decarboxylase [Lactococcus lactis]	78	68	174
450	1	1914	988	gi 1030068	[NAD(P)H oxidoreductase, isoflavone reductase homologue [Solanum tuberosum]	78	63	927
558	1	762	562	gi 1511588	bifunctional protein [Methanococcus jannaschii]	78	60	201
670	3	1152	1589	gi 1122759	unknown [Bacillus subtilis]	78	64	438
714	1	64	732	gi 143460	[37 kd minor sigma factor (rpoF, sigB; ttg start codon) [Bacillus ubtilis]	78	57	669
814	1	3	368	gi 1377833	unknown [Bacillus subtilis]	78	59	366
981	1	1381	692	gi 143802	GerC2 [Bacillus subtilis]	78	64	690
995	2	978	727	gi 296947	uridine kinase [Escherichia coli]	78	64	252
1045	1	3	401	gi 1407784	orf-1; novel antigen [Staphylococcus aureus]	78	61	399
1163	2	368	186	gi 410117	diaminopimelate decarboxylase [Bacillus subtilis]	78	54	183
2191	1	794	399	gi 215098	excisionase [Bacteriophage 154a]	78	65	396
2933	1	2	181	gi 1204436	pyruvate formate-lyase [Haemophilus influenzae]	78	73	180
3041	2	129	317	gi 624632	GlT-L [Escherichia coli]	78	53	189
3581	1	105	401	gi 763186	[3-ketoacyl-coA thiolase [Saccharomyces cerevisiae]	78	55	297
3709	1	3	230	gi 460689	TVG [Thermoactinomyces vulgaris]	78	58	228
3974	1	528	265	gi 558839	unknown [Bacillus subtilis]	78	65	264
3980	1	3	401	gi 39956	IIIGlc [Bacillus subtilis]	78	62	399
4056	1	647	354	gi 1256635	[dihydroxy-acid dehydratase [Bacillus subtilis]	78	55	294
4114	1	630	316	pir S09372 S093	hypothetical protein - Trypanosoma brucei	78	62	315
4185	1	3	179	gi 1339950	[large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	78	58	177

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4235	1	655	329	gi 558839	unknown [Bacillus subtilis]	78	60	327
4352	1	541	302	gi 603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis] gi 603768 HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis]	78	63	240
4368	1	612	307	gi 1353678	heavy-metal transporting P-type ATPase [Proteus mirabilis]	78	59	306
4461	1	428	216	gi 1276841	glutamate synthase (GOGAT) [Porphyra purpurea]	78	36	213
4530	1	474	238	gi 39956	IIGlc [Bacillus subtilis]	78	65	237
3	2	2969	2073	gi 1109684	Prov [Bacillus subtilis]	77	56	897
12	2	2426	1965	gi 467335	ribosomal protein L9 [Bacillus subtilis]	77	59	462
27	1	2	388	gi 1212728	YqhI [Bacillus subtilis]	77	63	387
39	2	590	1252	gi 40054	phenylalanyl-tRNA synthetase beta subunit (AA 1-804) [Bacillus btillis]	77	60	663
42	6	2704	2931	gi 606241	30S ribosomal subunit protein S14 [Escherichia coli] sp P02370 RS14_ECOLI 30S RIBOSOMAL PROTEIN S14. (SUB 2-101)	77	65	228
46	18	15459	16622	gi 297798	mitochondrial formate dehydrogenase precursor [Solanum tuberosum] pir JQ2272 JQ2272 formate dehydrogenase (EC 1.2.1.2) precursor, itochondrial - potato	77	55	1164
100	4	4562	4002	gi 1340128	ORF1 [Staphylococcus aureus]	77	54	561
102	8	5378	5713	gi 1311482	acetolactate synthase [Thermus aquaticus]	77	57	336
109	7	4742	5383	gi 710637	Unknown [Bacillus subtilis]	77	56	642
117	1	2	1228	gi 1237015	ORF4 [Bacillus subtilis]	77	53	1227
124	10	8323	7688	gi 405819	thymidine kinase [Bacillus subtilis]	77	63	636
147	3	1146	985	gi 849027	hypothetical 15.9-kDa protein [Bacillus subtilis]	77	37	162
152	10	7354	7953	gi 1205583	spermidine/putrescine transport ATP-binding protein [Haemophilus influenzae]	77	55	600
169	2	1004	1282	gi 473825	'elongation factor EF-Ts' [Escherichia coli]	77	58	279
184	2	380	1147	gi 1216314	esterase [Bacillus stearothermophilus]	77	60	768
189	7	3296	3868	gi 853809	ORF3 [Clostridium perfringens]	77	48	573
193	1	132	290	gi 1303788	YqeH [Bacillus subtilis]	77	54	159
195	8	8740	8414	gi 1499620	M. jannaschii predicted coding region MJ0798 [Methanococcus jannaschii]	77	44	327
205	8	5428	5204	gi 216340	ORF for adenylate kinase [Bacillus subtilis]	77	61	225

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
205	29	114795	114502	gi 786155	Ribosomal Protein L23 [Bacillus subtilis]	77	62	294
211	5	1908	2084	gi 410132	ORFX8 [Bacillus subtilis]	77	47	177
217	5	3478	4416	gi 496254	fibrinectin/fibrinogen-binding protein [Streptococcus pyogenes]	77	54	939
232	1	267	998	gi 11407784	orf-1; novel antigen [Staphylococcus aureus]	77	57	732
233	2	1819	1346	gi 467408	unknown [Bacillus subtilis]	77	61	474
243	3	2661	2299	gi 516155	unconventional myosin [Sus scrofa]	77	32	363
299	1	68	769	gi 467436	unknown [Bacillus subtilis]	77	54	702
301	4	1468	1283	gi 950071	ATP-bind. pyrimidine kinase [Mycoplasma capricolum] pir S48605 S48605	77	48	186
302	5	2741	3211	gi 508980	hypothetical protein - Mycoplasma capricolum SGC3 (fragment)			
302	7	3835	4863	gi 1147783	ruvB protein [Escherichia coli]	77	57	471
307	9	5402	4797	gi 1070015	protein-dependent [Bacillus subtilis]	77	60	606
312	1	99	1391	gi 143165	malic enzyme (EC 1.1.1.38) [Bacillus stearothermophilus] pir A33307 DEBSXS	77	62	1293
					malate dehydrogenase oxaloacetate-decarboxylating) (EC 1.1.1.38) -			
					Bacillus tearothermophilus			
312	2	1541	2443	gi 1399855	carboxyltransferase beta subunit [Synechococcus PCC7942]	77	58	903
321	5	5666	4596	gi 39844	fumarase (citG) (aa 1-462) [Bacillus subtilis]	77	65	1071
354	1	47	568	gi 1154634	YmaB [Bacillus subtilis]	77	57	522
365	1	2	1021	gi 143374	phosphoribosyl glycnamide synthetase (PUR-D; gtg start codon) Bacillus subtilis]	77	62	1020
374	1	1	708	gi 1405446	transketolase [Bacillus subtilis]	77	61	708
385	1	1128	565	gi 533099	endonuclease III [Bacillus subtilis]	77	63	564
392	2	594	1940	gi 556014	UDP-N-acetyl muramate-alanine ligase [Bacillus subtilis]	77	65	1347
					sp P40778 MURC BACSU UDP-N-ACETYLMURAMATE--ALANINE LIGASE (EC .3.2.8)			
					(UDP-N- ACETYLMURANOYL-L-ALANINE SYNTHETASE) (FRAGMENT).			
405	5	4079	3570	gi 1303912	YqhW [Bacillus subtilis]	77	64	510
487	4	1302	1472	gi 432427	ORF1 gene product [Acinetobacter calcoaceticus]	77	48	171
522	1	2	562	pir A01179 SYBS	tyrosine--tRNA ligase (EC 6.1.1.1) - Bacillus stearothermophilus	77	63	561

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
523	2	1587	1351	gi11387979	44% identity over 302 residues with hypothetical protein from <i>Synechocystis</i> sp, accession D64006 CD; expression induced by environmental stress; some similarity to glycosyl transferases; two potential membrane-spanning helices [Bacillus subtilis]	77	48	237
536	2	983	612	gi11433366	adenylosuccinate lyase (PUR-B) [Bacillus subtilis] pir C29326 WZBSDS	77	61	372
548	2	339	872	gi11433387	aspartate transcarbamylase [Bacillus subtilis]	77	56	534
597	1	2	481	gi1904198	hypothetical protein [Bacillus subtilis]	77	33	480
633	2	1747	1313	gi1387577	ORF1A [Bacillus subtilis]	77	64	435
642	1	85	360	gi146971	epiP gene product [Staphylococcus epidermidis]	77	61	276
659	1	125	1219	gi11072381	glutamyl-aminopeptidase [Lactococcus lactis]	77	62	1095
670	4	1587	1820	gi11122760	unknown [Bacillus subtilis]	77	58	234
789	1	2	391	gi11377823	aminopeptidase [Bacillus subtilis]	77	65	390
815	1	10	573	gi11303861	YqgN [Bacillus subtilis]	77	49	564
899	1	1	225	gi11204844	H. influenzae predicted coding region HI0594 [Haemophilus influenzae]	77	55	225
1083	1	3	188	gi1460828	B969 [Saccharomyces cerevisiae]	77	66	186
1942	1	415	209	gi1160047	p101/acidic basic repeat antigen [Plasmodium falciparum] pir A29232 A29232	77	38	207
2559	1	1	171	gi11499034	M. jannaschii predicted coding region MJ0255 [Methanococcus jannaschii]	77	61	171
2933	2	243	401	gi142370	pyruvate formate-lyase (AA 1-760) [Escherichia coli] ir S01788 S01788	77	72	159
2966	1	56	292	gi11524397	glycine betaine transporter OpuD [Bacillus subtilis]	77	45	237
2976	1	614	309	gi140003	oxoglutarate dehydrogenase (NADP+) [Bacillus subtilis] p P23129 ODO1_BACSU	77	60	306
2979	2	678	400	gi11204354	spore germination and vegetative growth protein [Haemophilus influenzae]	77	61	279
2988	1	601	377	gi1438465	Probable operon with orfF. Possible alternative initiation codon, ases	77	55	225
2990	1	331	167	gi1142562	ATP synthase epsilon subunit [Bacillus megaterium] pir B28599 PWBSEM H+-transporting ATP synthase (EC 3.6.1.34) psilon chain - Bacillus megaterium	77	63	165
3032	1	3	389	gi1488430	alcohol dehydrogenase 2 [Entamoeba histolytica]	77	56	387
3057	1	1	195	gi1468764	mocR gene product [Rhizobium meliloti]	77	50	195

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4008	1	726	400	gi 603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis]	77	52	327
				gi 603768	HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis]			
4048	1	703	386	gi 216278	gramicidin S synthetase 1 [Bacillus brevis]	77	55	318
4110	1	3	368	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	77	61	366
4115	1	1	348	gi 517205	167 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	77	65	348
4225	1	590	297	gi 1322245	mevalonate pyrophosphate decarboxylase [Rattus norvegicus]	77	60	294
4611	2	494	327	gi 508979	GTP-binding protein [Bacillus subtilis]	77	57	168
4668	1	361	182	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	77	61	180
25	1	2	1627	gi 1150620	MmsA [Streptococcus pneumoniae]	76	58	1626
38	5	1488	2537	pir A43577 A435	regulatory protein pfor - Clostridium perfringens	76	57	1050
52	5	2962	4041	gi 1161061	dioxygenase [Methylobacterium extorquens]	76	62	1080
56	120	127389	127955	gi 467402	unknown [Bacillus subtilis]	76	56	567
57	15	12046	12219	gi 1206040	weak similarity to keratin [Caenorhabditis elegans]	76	40	174
91	2	1062	2261	gi 475715	acetyl coenzyme A acetyltransferase (thiolase) [Clostridium cetobutylicum]	76	57	1200
98	2	818	1624	gi 467422	unknown [Bacillus subtilis]	76	62	807
98	5	2965	3228	gi 897793	y98 gene product [Pediococcus acidilactici]	76	52	264
98	8	5922	6326	gi 467427	methionyl-tRNA synthetase [Bacillus subtilis]	76	53	405
104	3	1322	1885	gi 216151	DNA polymerase (gene L; ttg start codon) [Bacteriophage SPO2] gi 579197	76	63	564
				SP02 DNA polymerase (aa 1-648) [Bacteriophage SPO2] pir A21498 DJBPS2 DNA-				
				directed DNA polymerase (EC 2.7.7.7) - phage PO2				
124	9	8134	7055	gi 853776	peptide chain release factor 1 [Bacillus subtilis] pir S55437 S55437	76	58	1080
				peptide chain release factor 1 - Bacillus ubtilis				
164	5	2832	3311	gi 1204976	prolyl-tRNA synthetase [Haemophilus influenzae]	76	53	480
168	2	2617	1841	gi 1177253	putative ATP-binding protein of ABC-type [Bacillus subtilis]	76	58	777
189	2	163	888	gi 467384	unknown [Bacillus subtilis]	76	63	726
235	3	2253	3518	gi 142936	folyl-polyglutamate synthetase [Bacillus subtilis] pir B40646 B40646 folC - Bacillus subtilis	76	53	1266
236	1	335	925	gi 1146197	putative [Bacillus subtilis]	76	54	591
237	8	5323	5541	gi 1279261	F13G3.6 [Caenorhabditis elegans]	76	47	219

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
263	5	5490	4585	gi 1510348	dihydrodipicolinate synthase [Methanococcus jannaschii]	76	49	906
304	3	1051	1794	gi 666982	putative membrane spanning subunit [Bacillus subtilis] pir S52382 S52382 probable membrane spanning protein - Bacillus ubtilis	76	60	744
312	4	3611	4624	gi 1143312	6-phospho-1-fructokinase (gtg start codon; EC 2.7.1.11) [Bacillus tearothermophilus]	76	56	1014
343	1	2	1036	gi 405956	yeeE [Escherichia coli]	76	59	1035
347	1	409	1701	gi 396304	acetylornithine deacetylase [Escherichia coli]	76	72	1293
358	1	672	1907	gi 1146215	39.0% identity to the Escherichia coli S1 ribosomal protein; putative [Bacillus subtilis]	76	58	1236
371	1	1	222	gi 537084	alternate gene name mgt; CG Site No. 497 [Escherichia coli] pir S56468 S56468 mgtA protein - Escherichia coli	76	61	222
379	4	4331	4858	gi 143268	dihydroliipoamide transsuccinylase (odhB; EC 2.3.1.61) [Bacillus ubtilis]	76	61	528
404	5	4022	4492	gi 1303823	YqfG [Bacillus subtilis]	76	60	471
411	1	2	307	gi 486025	ORF YKL027w [Saccharomyces cerevisiae]	76	55	306
472	3	4356	2854	gi 1405464	Alst [Bacillus subtilis]	76	57	1503
546	1	273	995	gi 153821	streptococcal pyrogenic exotoxin type C (speC) precursor Streptococcus pyogenes]	76	36	723
588	1	1054	557	gi 1002520	MutS [Bacillus subtilis]	76	61	498
591	1	16	735	gi 885934	ClpB [Synechococcus sp.]	76	44	720
602	2	175	798	gi 1486422	OppD homologue [Rhizobium sp.]	76	52	624
619	2	547	290	gi 330613	major capsid protein [Human cytomegalovirus]	76	47	258
660	4	2568	3302	gi 904199	hypothetical protein [Bacillus subtilis]	76	55	735
677	1	452	228	gi 40177	spoOF gene product [Bacillus subtilis]	76	58	225
962	1	24	206	gi 1142443	adenylosuccinate synthetase [Bacillus subtilis] sp P29726 PURA_BACSU ADENYLOSUCCINATE SYNTHETASE (EC 6.3.4.4) IMP--ASPARTATE LIGASE).	76	67	183
978	1	1158	580	gi 1511333	M. jannaschii predicted coding region MJ1322 [Methanococcus jannaschii]	76	56	579
997	1	486	244	gi 467154	No definition line found [Mycobacterium leprae]	76	38	243
1563	1	529	266	gi 1303984	YqkG [Bacillus subtilis]	76	52	264
2184	1	361	182	gi 506706	CapJ [Staphylococcus aureus]	76	38	180
2572	1	1	387	gi 153898	transport protein [Salmonella typhimurium]	76	65	387

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
2942	1	29	400	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	76	59	372
2957	1	377	216	gi 1511251	hypothetical protein (SP:P42404) [Methanococcus jannaschii]	76	47	162
2980	1	554	279	gi 1405464	Alst [Bacillus subtilis]	76	53	276
3015	1	649	326	gi 408115	ornithine acetyltransferase [Bacillus subtilis]	76	61	324
3124	1	13	174	gi 882705	ORF_o401 [Escherichia coli]	76	65	162
3179	1	3	161	gi 168477	ferredoxin-dependent glutamate synthase [Zea mays] pir A38596 A38596 glutamate synthase (ferredoxin) (EC 1.4.7.1) - aize	76	53	159
3789	1	2	379	gi 39956	IIIGlc [Bacillus subtilis]	76	55	378
3892	1	3	314	gi 1510398	ferripyochelin binding protein [Methanococcus jannaschii]	76	52	312
3928	1	798	400	gi 143016	permease [Bacillus subtilis]	76	59	399
4159	1	757	386	sp P80544 MRSP_	METHICILLIN-RESISTANT SURFACE PROTEIN (FRAGMENTS).	76	66	372
4204	1	17	331	gi 296464	ATPase [Lactococcus lactis]	76	56	315
4398	1	494	249	gi 987255	Menkes disease gene [Homo sapiens]	76	48	246
4506	1	2	313	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	76	47	312
4546	1	477	247	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	76	61	231
4596	1	379	191	gi 560027	cellulose synthase [Acetobacter xylinum]	76	70	189
4	5	5257	4337	gi 882532	ORF_o294 [Escherichia coli]	75	59	921
6	1	164	952	gi 40960	OTCase [Escherichia coli]	75	56	789
12	3	5935	3944	gi 467336	unknown [Bacillus subtilis]	75	57	1992
23	18	18272	17310	gi 1296433	O-acetylserine sulphydrylase B [Alcaligenes eutrophus]	75	55	963
25	3	2356	3393	gi 1502419	PlsX [Bacillus subtilis]	75	56	1038
36	8	5765	6037	gi 1256517	unknown [Schizosaccharomyces pombe]	75	45	273
46	13	11186	12058	gi 48972	nitrate transporter [Synechococcus sp.]	75	46	873
51	7	3474	3677	gi 143607	sporulation protein [Bacillus subtilis]	75	61	204
53	16	16850	16590	gi 143402	recombination protein (ttg start codon) [Bacillus subtilis] gi 1303923 RecN [Bacillus subtilis]	75	51	261
74	3	3572	2568	gi 1204847	ornithine carbamoyltransferase [Haemophilus influenzae]	75	61	1005

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
85	3	4628	3930	gi1143368	phosphoribosylformyl glycinamide synthetase I (PUR-L; gtg start odon) [Bacillus subtilis]	75	63	699
85	5	5588	4878	gi1143367	phosphoribosyl aminoidazole succinocarboxamide synthetase (PUR-C; tg start codon) [Bacillus subtilis]	75	55	711
85	8	6625	7530	gi11303916	YqiA [Bacillus subtilis]	75	53	906
87	3	2340	3590	gi11064813	homologous to sp:PHOR_BACSU [Bacillus subtilis]	75	56	1251
87	6	6084	6896	gi11064810	function unknown [Bacillus subtilis]	75	61	813
108	2	1844	1503	gi11001824	hypothetical protein [Synechocystis sp.]	75	51	342
110	3	1748	3727	gi11147593	putative ppGpp synthetase [Streptomyces coelicolor]	75	55	1980
110	7	4353	5252	gi11177251	clwD gene product [Bacillus subtilis]	75	75	900
120	14	11266	10649	gi11524394	ORF-2 upstream of gbsAB operon [Bacillus subtilis]	75	55	618
121	5	2050	4221	gi11154632	NrdE [Bacillus subtilis]	75	54	2172
124	1	283	143	gi1405622	unknown [Bacillus subtilis]	75	56	141
128	1	81	1139	gi1143316	[gap] gene products [Bacillus megaterium]	75	48	1059
130	8	5760	5903	gi11256654	54.8% identity with Neisseria gonorrhoeae regulatory protein PilB; putative [Bacillus subtilis]	75	62	144
136	2	4480	3185	gi1467403	seryl-tRNA synthetase [Bacillus subtilis]	75	54	1296
161	10	5439	5798	gi11001195	hypothetical protein [Synechocystis sp.]	75	55	360
172	4	3819	2995	gi1755153	ATP-binding protein [Bacillus subtilis]	75	52	825
179	1	2024	1107	gi1143037	porphobilinogen deaminase [Bacillus subtilis]	75	58	918
195	10	9529	9374	sp125745 YCFB_	HYPOTHEICAL PROTEIN IN PURB 5'REGION (ORF-15) (FRAGMENT).	75	60	156
200	4	2605	4596	gi1142440	ATP-dependent nuclease [Bacillus subtilis]	75	56	1992
206	3	6900	5620	gi11256135	YbbF [Bacillus subtilis]	75	53	1281
216	2	159	389	gi11052800	unknown [Schizosaccharomyces pombe]	75	58	231
229	1	29	847	gi11205958	branched chain aa transport system II carrier protein [Haemophilus influenzae]	75	49	819
230	2	518	1714	gi1971337	nitrite extrusion protein [Bacillus subtilis]	75	53	1197
231	1	2240	1122	gi11002521	MutL [Bacillus subtilis]	75	54	1119
233	3	1314	1859	gi1467405	unknown [Bacillus subtilis]	75	59	546

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
269	1	325	164	gi11511246	{methyl coenzyme M reductase system, component A2 [Methanococcus jannaschii]}	75	50	162
292	1	1389	772	gi11511604	{M. jannaschii predicted coding region MJ1651 [Methanococcus jannaschii]}	75	46	618
304	4	1773	2261	gi11205328	{surfactin [Haemophilus influenzae]}	75	55	489
312	3	2437	3387	gi1285621	{undefined open reading frame [Bacillus stearothermophilus]}	75	62	951
312	5	4622	6403	gi11041097	{Pyruvate Kinase [Bacillus psychrophilus]}	75	57	1782
319	1	353	877	gi11212728	{YqhI [Bacillus subtilis]}	75	54	525
320	5	4321	5031	gi11070361	{OMP decarboxylase [Lactococcus lactis]}	75	56	711
320	6	5010	5642	gi1143394	{OMP-PRPP transferase [Bacillus subtilis]}	75	60	633
337	4	1519	2088	gi1487433	{citrate synthase II [Bacillus subtilis]}	75	58	570
394	2	669	1271	gi1304976	{matches PS00017: ATP_GTP_A and PS00301: ERFATOR_GTP; similar to longation factor G, TetM/TetO tetracycline-resistance proteins Escherichia coli}	75	51	603
423	1	127	570	gi11183839	{unknown [Pseudomonas aeruginosa]}	75	59	444
433	2	1603	1929	gi1149211	{acetolactate synthase [Klebsiella pneumoniae]}	75	63	327
446	2	176	1540	gi1312441	{dihydroorotase [Bacillus caldolyticus]}	75	62	1365
486	1	494	249	gi11149682	{potF gene product [Clostridium perfringens]}	75	55	246
496	1	3	794	gi1143582	{spoIIIEA protein [Bacillus subtilis]}	75	59	792
498	2	824	1504	gi1143328	{phoP protein (put.); putative [Bacillus subtilis]}	75	47	681
499	2	1061	1624	gi11387979	{44% identity over 302 residues with hypothetical protein from Synechocystis sp, accession D64006_CD; expression induced by environmental stress; some similarity to glycosyl transferases; two potential membrane-spanning helices [Bacillus subtilis]}	75	51	564
568	1	641	453	pir JC4110 JC41	{triacylglycerol lipase (EC 3.1.1.3) 2 - Mycoplasma mycoides subsp. mycoides (SGC3)}	75	50	189
613	2	430	233	gi1330993	{tegument protein [Saimiriine herpesvirus 2]}	75	75	198
621	1	1	525	gi1529754	{spec [Streptococcus pyogenes]}	75	43	525
642	5	1809	2474	gi11176401	{EpiG [Staphylococcus epidermidis]}	75	51	666
646	2	454	657	gi1172442	{ribonuclease P [Saccharomyces cerevisiae]}	75	37	204
657	1	3	347	gi1882541	{ORF_o236 [Escherichia coli]}	75	47	345
750	1	1662	832	gi146971	{epiP gene product [Staphylococcus epidermidis]}	75	57	831

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
754	1	2	481	gi 1303901	YqHT [Bacillus subtilis]	75	57	480
763	2	563	393	gi 1205145	multidrug resistance protein [Haemophilus influenzae]	75	51	171
775	1	961	482	pir B36889 B368	leuA protein, inactive - Lactococcus lactis subsp. lactis (strain IL1403)	75	63	480
793	1	1	180	gi 143316	[gap] gene products [Bacillus megaterium]	75	57	180
800	1	318	160	gi 509411	NERA protein [Azorhizobium caulinodans]	75	34	159
811	1	1117	560	gi 143434	Rho Factor [Bacillus subtilis]	75	60	558
940	1	493	329	gi 1276985	arginase [Bacillus caldovelox]	75	50	165
971	2	37	252	gi 1001373	hypothetical protein [Synechocystis sp.]	75	58	216
1059	1	384	232	gi 726480	L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus ubtilis]	75	67	153
1109	2	219	374	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis]	75	53	156
				pir A27650 A27650	regulatory protein phoR - Bacillus subtilis			
				sp P23545 PHOR_BACSU	ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOR (EC 2.7.3.-)			
1268	1	271	137	gi 304135	ornithine acetyltransferase [Bacillus stearothermophilus]	75	63	135
				sp Q07908 ARGJ_BACST	GLUTAMATE N-ACETYLTRANSFERASE (EC 2.3.1.35) ORNITHINE ACETYLTRANSFERASE) (ORNITHINE TRANSACETYLASE) (OATASE) / MINO-ACID ACETYLTRANSFERASE (EC 2.3.1.1) (N-ACETYLGLUTAMATE YNTHA			
1500	1	324	163	gi 1205488	excinuclease ABC subunit B [Haemophilus influenzae]	75	57	162
1529	1	798	400	gi 1002521	MutL [Bacillus subtilis]	75	54	399
3010	1	770	387	gi 1204435	pyruvate formate-lyase activating enzyme [Haemophilus influenzae]	75	54	384
3105	1	1	180	gi 1041097	Pyruvate Kinase [Bacillus psychrophilus]	75	57	180
3117	1	45	212	gi 899317	peptide synthetase module [Microcystis aeruginosa] pir S49111 S49111	75	42	168
					probable amino acid activating domain - icrocystis aeruginosa (fragment)			
					(SUB 144-528)			
3139	2	139	345	gi 145294	adenine phosphoribosyl-transferase [Escherichia coli]	75	66	207
3880	1	618	310	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	75	58	309
3911	1	48	401	gi 433991	ATP synthase subunit beta [Bacillus subtilis]	75	68	354
3957	1	2	379	pir D36889 D368	3-isopropylmalate dehydratase (EC 4.2.1.33) chain leuC - Lactococcus lactis subsp. lactis (strain IL1403)	75	65	378
4005	1	5	259	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	75	48	255
4080	1	73	333	gi 415855	deoxyribose aldolase [Mycoplasma hominis]	75	59	261

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4111	1	1	339	gi 1149435	putative [Lactococcus lactis]	75	57	339
4136	1	602	303	gi 450688	hsdM gene of EcoprrI gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	75	56	300
4144	1	668	336	gi 148972	nitrate transporter [Synechococcus sp.]	75	49	333
4237	1	664	374	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	75	55	291
4306	2	73	318	gi 294260	major surface glycoprotein [Pneumocystis carinii]	75	68	246
4343	1	715	359	gi 1204652	methylated-DNA--protein-cysteine methyltransferase [Haemophilus influenzae]	75	52	357
4552	1	620	312	gi 296464	ATPase [Lactococcus lactis]	75	55	309
38	9	5776	6126	gi 1443793	NupC [Escherichia coli]	74	50	351
50	8	6910	6221	gi 1239988	hypothetical protein [Bacillus subtilis]	74	55	690
56	9	10770	12221	gi 1000451	TreP [Bacillus subtilis]	74	57	1452
64	2	2266	1622	gi 141015	aspartate-tRNA ligase [Escherichia coli]	74	57	645
66	6	5063	4848	gi 1212729	YqhJ [Bacillus subtilis]	74	47	216
67	18	14334	14897	gi 1510631	endoglucanase [Methanococcus jannaschii]	74	52	564
102	15	12561	13136	gi 1149429	putative [Lactococcus lactis]	74	67	576
102	16	13121	14419	gi 1149435	putative [Lactococcus lactis]	74	57	1299
108	4	4873	3902	gi 39478	ATP binding protein of transport ATPases [Bacillus firmus] ir S15486 S15486 ATP-binding protein - Bacillus firmus p P26946 YATR_BACFI HYPOTHETICAL ATP-BINDING TRANSPORT PROTEIN.	74	59	972
116	5	8574	7093	gi 1205430	dipeptide transport system permease protein [Haemophilus influenzae]	74	49	1482
120	7	4342	4803	gi 1146970	ribonucleoside triphosphate reductase [Escherichia coli] pir A47331 A47331 anaerobic ribonucleotide reductase - Escherichia oli	74	58	462
121	7	5961	6581	gi 1107528	ttg start [Campylobacter coli]	74	51	621
128	3	2320	3531	gi 1143318	phosphoglycerate kinase [Bacillus megaterium]	74	57	1212
130	7	5237	5791	gi 1256653	DNA-binding protein [Bacillus subtilis]	74	60	555
136	3	6745	5150	gi 1143076	histidase [Bacillus subtilis]	74	58	1596
145	2	664	1368	gi 1407773	devA gene product [Anabaena sp.]	74	45	705
152	1	552	277	gi 1377833	unknown [Bacillus subtilis]	74	54	276

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
164	10	11064	11375	gi1580900	ORF3 gene product [Bacillus subtilis]	74	52	312
175	2	3109	2624	gi1642656	unknown [Rhizobium meliloti]	74	34	486
175	9	6064	5612	gi1854656	Na/H antiporter system ORF2 [Bacillus alcalophilus]	74	46	453
195	11	11346	10339	gi11204430	hypothetical protein (SP:P25745) [Haemophilus influenzae]	74	55	1008
205	17	9619	9059	gi11044979	ribosomal protein L6 [Bacillus subtilis]	74	64	561
236	7	5574	6710	gi11146207	putative [Bacillus subtilis]	74	63	1137
241	3	4521	3334	gi1694121	malate thiokinase [Methylobacterium extorquens]	74	52	1188
246	6	3305	2799	gi1467374	single strand DNA binding protein [Bacillus subtilis]	74	64	507
249	4	6551	5313	gi11524397	glycine betaine transporter OpuD [Bacillus subtilis]	74	55	1239
261	7	4389	4081	gi1809542	ChrB protein [Erwinia chrysanthemi]	74	42	309
278	6	5714	4665	gi11204872	ATP-binding protein [Haemophilus influenzae]	74	54	1050
309	1	1220	666	gi11205579	hypothetical protein (GB:U14003_302) [Haemophilus influenzae]	74	53	555
315	2	1473	862	gi1143398	quinol oxidase [Bacillus subtilis]	74	57	612
320	1	1	1065	gi1143389	glutaminase of carbamyl phosphate synthetase [Bacillus subtilis]	74	60	1065
				pir E39845 E39845 carbamoyl-phosphate synthase glutamine-hydrolyzing) (EC 6.3.5.5), pyrimidine-repressible, small hain - Bacillus subtilis				
380	2	382	1128	gi1534857	ATPase subunit a [Bacillus stearothermophilus]	74	56	747
405	2	1742	1311	gi11303915	YqhZ [Bacillus subtilis]	74	65	432
433	5	2503	3270	gi1473902	alpha-acetolactate synthase [Lactococcus lactis]	74	56	768
452	1	1	942	gi1413982	ipa-58r gene product [Bacillus subtilis]	74	52	942
461	1	3	1193	gi1558494	homoserine dehydrogenase [Bacillus subtilis]	74	51	1191
461	2	1174	1407	gi140211	threonine synthase (thrC) (AA 1-352) [Bacillus subtilis] ir A25364 A25364	74	56	234
				threonine synthase (EC 4.2.99.2) - Bacillus btillis				
462	2	402	734	gi1142520	thioredoxin [Bacillus subtilis]	74	62	333
478	1	574	320	gi11499005	glycyl-tRNA synthetase [Methanococcus jannaschii]	74	52	255
501	2	739	1740	gi1217040	acid glycoprotein [Streptococcus pyogenes]	74	58	1002
551	2	4083	2791	gi1143040	glutamate-1-semialdehyde 2,1-aminotransferase [Bacillus subtilis]	74	51	1293
				pir D42728 D42728 glutamate-1-semialdehyde 2,1-aminomutase (EC .4.3.8) - Bacillus subtilis				

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
573	1	1	477	gi 1006605	hypothetical protein [Synechocystis sp.]	74	45	477
596	2	1780	1298	gi 1303853	YqgF [Bacillus subtilis]	74	55	483
618	2	2924	1758	gi 1146237	21.4% of identity to trans-acting transcription factor of Sacharomyces cerevisiae; 25% of identity to sucrose synthase of Zea mays; putative [Bacillus subtilis]	74	55	1167
659	2	1269	1595	gi 1072380	ORF3 [Lactococcus lactis]	74	62	327
724	1	373	188	gi 143374	phosphoribosyl glycinamide synthetase (PUR-D; gtg start codon) Bacillus subtilis]	74	58	186
743	2	604	1209	gi 153833	ORF1; putative [Streptococcus parasanguis]	74	50	606
836	1	2	259	gi 143458	ORF V [Bacillus subtilis]	74	47	258
989	2	443	724	gi 1303994	YqkM [Bacillus subtilis]	74	46	282
1106	1	1	492	gi 46970	epiD gene product [Staphylococcus epidermidis]	74	54	492
1135	2	373	528	gi 413948	ipa-24d gene product [Bacillus subtilis]	74	48	156
1234	1	817	452	gi 495245	recJ gene product [Erwinia chrysanthemi]	74	36	366
2586	1	2	238	gi 1149701	sbcc gene product [Clostridium perfringens]	74	62	237
2959	1	798	400	gi 1405454	aconitase [Bacillus subtilis]	74	60	399
2962	1	650	363	gi 450686	3-phosphoglycerate kinase [Thermotoga maritima]	74	58	288
2983	1	3	191	gi 1303893	YqhL [Bacillus subtilis]	74	56	189
3018	1	2	223	gi 143040	glutamate-1-semialdehyde 2,1-aminotransferase [Bacillus subtilis]	74	56	222
				pir D42728 D42728	glutamate-1-semialdehyde 2,1-aminomutase (EC .4.3.8) - Bacillus subtilis			
3038	1	510	256	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	74	57	255
3062	1	374	189	gi 1107528	ttg start [Campylobacter coli]	74	51	186
4035	1	184	360	gi 1022725	unknown [Staphylococcus haemolyticus]	74	64	177
4045	1	607	305	gi 1510977	M. jannaschii predicted coding region MJ0938 [Methanococcus jannaschii]	74	41	303
4283	1	471	304	gi 520844	orf4 [Bacillus subtilis]	74	58	168
4449	1	3	221	gi 580910	peptide-synthetase ORF1 [Bacillus subtilis]	74	54	219
4587	1	458	231	gi 1370207	orf6 [Lactobacillus sake]	74	59	228

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4603	1	29	214	gi 146208	glutamate synthase large subunit (EC 2.6.1.53) [Escherichia coli] pir A29617 A29617 glutamate synthase (NADPH) (EC 1.4.1.13) large hain - Escherichia coli	74	60	186
4670	1	366	184	gi 1256135	YbbF [Bacillus subtilis]	74	61	183
5	10	7953	7162	gi 143727	[putative [Bacillus subtilis]	73	42	792
11	2	2454	1372	gi 166338	[dihydroorotate dehydrogenase [Agrocye aegerita]	73	55	1083
14	1	2024	1020	gi 143373	[phosphoribosyl aminimidazole carboxy formyl ornyltransferase/inosine monophosphate cyclohydrolase (PUR-H(J)) Bacillus subtilis]	73	54	1005
23	5	5426	4635	gi 1468939	[meso-2,3-butanediol dehydrogenase (D-acetoin forming) [Klebsiella pneumoniae]	73	58	792
23	17	117379	16360	gi 1297060	[ornithine cyclodeaminase [Rhizobium meliloti]	73	37	1020
29	2	692	1273	gi 467442	[stage V sporulation [Bacillus subtilis]	73	54	582
31	5	6467	4914	gi 414000	[ipa-76d gene product [Bacillus subtilis]	73	55	1554
37	8	8658	7402	gi 1429259	[pept gene product [Bacillus subtilis]	73	59	1257
37	9	7738	7562	gi 168367	[alpha-isopropylmalate isomerase (put.); putative [Rhizomucor ircinelloides]	73	52	177
38	7	3931	4896	gi 405885	[yeiN [Escherichia coli]	73	58	966
44	6	5041	4238	gi 580895	[unknown [Bacillus subtilis]	73	53	804
44	11	7767	8306	gi 42009	[moaB gene product [Escherichia coli]	73	50	540
45	3	2439	3080	gi 1109685	[ProW [Bacillus subtilis]	73	47	642
54	13	14036	13794	gi 413931	[ipa-7d gene product [Bacillus subtilis]	73	61	243
59	4	1430	2248	gi 147923	[threonine dehydratase 2 (EC 4.2.1.16) [Escherichia coli]	73	53	819
65	1	1458	730	gi 677944	[AppF [Bacillus subtilis]	73	56	729
80	2	1375	860	gi 580932	[murD gene product [Bacillus subtilis]	73	53	516
102	13	10124	11179	gi 580891	[3-isopropylmalate dehydrogenase (AA 1 - 365) [Bacillus subtilis] pir A26522 A26522 3-isopropylmalate dehydrogenase (EC 1.1.1.85) - acillus subtilis	73	55	1056
109	2	3493	2600	gi 1510849	[M. jannaschii predicted coding region MJ0775 [Methanococcus jannaschii]	73	40	894
120	8	4782	5756	gi 146970	[ribonucleoside triphosphate reductase [Escherichia coli] pir A47331 A47331 anaerobic ribonucleotide reductase - Escherichia oli	73	56	975
120	9	5726	6223	gi 1204333	[anaerobic ribonucleoside-triphosphate reductase [Haemophilus influenzae]	73	62	498

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
132	5	4151	4363	gi 871048	HPSR2 - heavy chain potential motor protein [Giardia intestinalis]	73	43	213
140	6	5952	4324	gi 634107	kdpB [Escherichia coli]	73	59	1629
142	6	7060	5939	gi 410125	ribG gene product [Bacillus subtilis]	73	57	1122
149	4	1866	1717	gi 460892	heparin binding protein-44, HBP-44 [mice, Peptide, 360 aa] pir JX0281 JX0281 heparin-binding protein-44 precursor - mouse gi 220434 ORF [Mus musculus] (SUB 2-360)	73	53	150
158	1	1	1431	gi 882504	ORF_f560 [Escherichia coli]	73	57	1431
174	6	5352	4525	gi 1146240	ketopantoate hydroxymethyltransferase [Bacillus subtilis]	73	55	828
175	8	5537	5178	gi 854657	Na/H antiporter system ORF3 [Bacillus alcalophilus]	73	56	360
186	5	6593	5493	gi 467477	unknown [Bacillus subtilis]	73	48	1101
249	6	6283	5729	gi 1524397	glycine betaine transporter OpuD [Bacillus subtilis]	73	56	555
265	4	1873	2280	gi 39848	U3 [Bacillus subtilis]	73	41	408
270	1	328	582	gi 780461	1220 kDa polyprotein [African swine fever virus]	73	53	255
278	4	4283	3618	gi 1208965	hypothetical 23.3 kd protein [Escherichia coli]	73	49	666
279	3	4984	3593	gi 1185288	isochorismate synthase [Bacillus subtilis]	73	58	1392
291	4	1207	1575	gi 1511440	glutamine--fructose-6-phosphate transaminase [Methanococcus jannaschii]	73	63	369
299	2	735	1166	gi 467437	unknown [Bacillus subtilis]	73	58	432
299	5	2050	3234	gi 467439	temperature sensitive cell division [Bacillus subtilis]	73	53	1185
334	1	1237	728	gi 536655	ORF_YBR244w [Saccharomyces cerevisiae]	73	43	510
336	2	1827	1036	gi 790943	urea amidolyase [Bacillus subtilis]	73	51	792
374	3	1389	1874	gi 1405451	YneJ [Bacillus subtilis]	73	55	486
433	4	1916	2554	gi 473902	alpha-acetolactate synthase [Lactococcus lactis]	73	54	639
509	2	1795	1028	gi 467483	unknown [Bacillus subtilis]	73	56	768
513	1	1709	918	gi 1146220	INAD+ dependent glycerol-3-phosphate dehydrogenase [Bacillus subtilis]	73	56	792
533	2	239	733	gi 1510605	hypothetical protein (SP:P42297) [Methanococcus jannaschii]	73	44	495
546	2	1148	2815	gi 41748	hscM protein (AA 1-520) [Escherichia coli]	73	52	1668
549	1	762	382	gi 1314847	CinA [Bacillus subtilis]	73	57	381
567	1	1346	675	gi 410137	ORFX13 [Bacillus subtilis]	73	58	672

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
716	2	654	1112	gi11256623	exodeoxyribonuclease [Bacillus subtilis]	73	56	459
772	1	3	677	gi1142010	Shows 70.2% similarity and 48.6% identity to the EnvM protein of almonella typhimurium [Anabaena sp.]	73	57	675
774	1	3	209	gi1409286	bmrU [Bacillus subtilis]	73	52	207
782	1	1	402	gi1143320	[gap] gene products [Bacillus megaterium]	73	56	402
789	2	451	762	gi11063246	low homology to P14 protein of Heamophilus influenzae and 14.2 kDa protein of Escherichia coli [Bacillus subtilis]	73	56	312
796	1	3	911	gi1853754	ABC transporter [Bacillus subtilis]	73	58	909
806	3	1209	949	gi1143786	tryptophanyl-tRNA synthetase (EC 6.1.1.2) [Bacillus subtilis] pirJT0481 YWB5 tryptophan--tRNA ligase (EC 6.1.1.2) - Bacillus ubtilis	73	51	261
816	2	4839	3097	gi141748	hsdM protein (AA 1-520) [Escherichia coli]	73	52	1743
839	1	798	400	gi1886906	argininosuccinate synthetase [Streptomyces clavuligerus] pirIS57659 S57659 argininosuccinate synthase (EC 6.3.4.5) - treptomyces clavuligerus	73	59	399
857	1	3	290	gi1348052	acetoin utilization protein [Bacillus subtilis]	73	50	288
1008	1	790	398	gi140100	rodC (tag3) polypeptide (AA 1-746) [Bacillus subtilis] irIS06049 S06049 rodC protein - Bacillus subtilis p P13485 TAGF_BACSU TECHNOIC ACID BIOSYNTHESIS PROTEIN F.	73	41	393
1018	1	1	213	gi1529357	No definition line found [Caenorhabditis elegans] sp P46975 STT3_CAEEL OLIGOSACCHARYL TRANSFERASE STT3 SUBUNIT ONOLOG.	73	53	213
1033	1	3	491	gi1142706	comG1 gene product [Bacillus subtilis]	73	51	489
1174	1	395	204	gi11149513	alpha3a subunit of laminin 5 [Homo sapiens]	73	60	192
1175	1	655	329	gi1473817	'ORF' [Escherichia coli]	73	57	327
1187	1	3	209	gi1580870	ipa-37d qoxA gene product [Bacillus subtilis]	73	52	207
1206	1	72	245	gi1144816	formyltetrahydrofolate synthetase (FTHFS) (ttg start codon) (EC .3.4.3) [Moorella thermoacetica]	73	43	174
1454	1	423	241	gi11213253	unknown [Schizosaccharomyces pombe]	73	53	183
1469	1	517	260	gi11303787	Yqeg [Bacillus subtilis]	73	55	258
1761	1	374	189	gi19135	Mst26Aa gene product [Drosophila simulans]	73	34	186
1849	1	467	243	gi1162307	DNA topoisomerase II [Trypanosoma cruzi]	73	60	225
2055	1	2	400	gi1559381	P47K protein [Rhodococcus erythropolis]	73	34	399
2556	1	2	244	gi1145925	fecB [Escherichia coli]	73	62	243

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
2947	2	549	400	gi 1184680	polynucleotide phosphorylase [Bacillus subtilis]	73	51	150
2956	1	746	375	gi 143397	quinol oxidase [Bacillus subtilis]	73	58	372
3037	1	655	329	gi 143091	acetolactate synthase [Bacillus subtilis]	73	55	327
3115	1	385	194	gi 323866	overlapping out-of-phase protein [Eggplant mosaic virus]	73	53	192
				sp P20129 V70K_EPMV 70 KD PROTEIN.				
3603	2	700	527	gi 11439521	glutaryl-CoA dehydrogenase precursor [Mus musculus]	73	48	174
3743	1	798	400	gi 450688	hsdM gene of Ecoprr1 gene product [Escherichia coli] pir S38437 S38437 hsdM	73	54	399
				protein - Escherichia coli pir S09629 S09629 hypothetical protein A -				
				Escherichia coli {SUB 40-520}				
3752	1	640	359	gi 1524193	unknown [Mycobacterium tuberculosis]	73	59	282
3852	1	2	181	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	73	68	180
3914	1	475	239	pir S13490 S134	Hydroxymethylglutaryl-CoA synthase (EC 4.1.3.5) - Chicken (fragment)	73	53	237
3914	2	570	343	gi 528991	unknown [Bacillus subtilis]	73	38	228
4069	1	2	316	gi 40003	oxoglutarate dehydrogenase (NADP+) [Bacillus subtilis] p P23129 ODO1_BACSU	73	55	315
				2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (EC 2.4.2) (ALPHA- KETOGLUTARATE				
				DEHYDROGENASE).				
4165	1	715	365	gi 1439521	glutaryl-CoA dehydrogenase precursor [Mus musculus]	73	48	351
4196	1	1	177	gi 809660	deoxyribose-phosphate aldolase [Bacillus subtilis] pir S49455 S49455	73	60	177
				deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis				
4202	1	572	378	gi 528991	unknown [Bacillus subtilis]	73	38	195
4314	1	2	193	gi 436797	N-acyl-L-amino acid amidohydrolase [Bacillus stearothermophilus]	73	47	192
				sp P37112 AMA_BACST N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC .5.1.14)				
				(AMINOACYLASE).				
4393	1	3	263	gi 216267	ORE2 [Bacillus megaterium]	73	47	261
35	2	903	1973	gi 1146196	phosphoglycerate dehydrogenase [Bacillus subtilis]	72	53	1071
38	22	19094	17877	gi 602031	similar to trimethylamine DH [Mycoplasma capricolum] pir S49950 S49950	72	54	1218
				probable trimethylamine dehydrogenase (EC .5.99.7) - Mycoplasma capricolum				
				(SGC3) (fragment)				
38	23	18134	19162	gi 413968	ipa-44d gene product [Bacillus subtilis]	72	54	1029
44	19	11895	12953	gi 516272	unknown [Bacillus subtilis]	72	49	1059
48	7	6248	7117	gi 43499	pyruvate synthase [Halobacterium halobium]	72	49	870
50	7	6563	5691	gi 1205399	proton glutamate symport protein [Haemophilus influenzae]	72	53	873

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
53	9	110521	9259	gi 1303956	YqjE [Bacillus subtilis]	72	52	1263
56	123	129549	12995	gi 467471	unknown [Bacillus subtilis]	72	47	447
69	4	5298	4123	gi 1354775	pfos/R [Treponema pallidum]	72	46	1176
69	5	4377	4982	gi 904198	hypothetical protein [Bacillus subtilis]	72	43	606
73	1	2	856	gi 142997	glycerol uptake facilitator [Bacillus subtilis]	72	59	855
98	113	9371	110258	gi 467435	unknown [Bacillus subtilis]	72	50	888
127	1	1	1593	gi 217144	alanine carrier protein [thermophilic bacterium PS3] pir A45111 A45111	72	56	1593
					alanine transport protein - thermophilic acterium PS-3			
131	1	5197	2600	gi 153952	polymerase III polymerase subunit (dnaE) [Salmonella typhimurium]	72	53	2598
					pir A45915 A45915 DNA-directed DNA polymerase (EC 2.7.7.7) III lpha chain			
					- Salmonella typhimurium			
141	4	1040	1978	gi 1405446	transketolase [Bacillus subtilis]	72	54	939
149	8	2819	2535	gi 606234	secY [Escherichia coli]	72	44	285
149	17	5472	5245	gi 1304472	DNA polymerase [Unidentified phycodnavirus clone OTU4]	72	55	228
154	1	1	210	gi 1205620	ferritin like protein [Haemophilus influenzae]	72	40	210
155	1	2207	1320	gi 391610	farnesyl diphosphate synthase [Bacillus stearothermophilus]	72	57	888
					pir JX0257 JX0257 geranyltransferase (EC 2.5.1.10) - Bacillus tearothermophilus			
180	1	2	328	gi 433630	A180 [Saccharomyces cerevisiae]	72	62	327
184	3	1145	3553	gi 1205110	virulence associated protein homolog [Haemophilus influenzae]	72	49	2409
195	2	1923	1279	gi 1001730	hypothetical protein [Synechocystis sp.]	72	45	645
206	13	114646	15869	gi 1064807	ORTHININE AMINOTRANSFERASE [Bacillus subtilis]	72	50	1224
209	2	462	932	gi 1204666	hypothetical protein (GB:X73124.53) [Haemophilus influenzae]	72	60	471
215	2	764	522	gi 881513	insulin receptor homolog [Drosophila melanogaster] pir S57245 S57245	72	63	243
					insulin receptor homolog - fruit fly (Drosophila elanogaster) {SUB 46-2146}			
224	1	2	790	gi 949974	sucrose repressor [Staphylococcus xylosus]	72	54	789
233	1	1526	765	gi 1408493	homologous to SwissProt:YIDA_ECOLI hypothetical protein [Bacillus subtilis]	72	52	762
240	1	220	1485	gi 537049	ORF_o470 [Escherichia coli]	72	52	1266
245	1	3	1340	gi 1204578	hypothetical protein (GB:U06949.1) [Haemophilus influenzae]	72	46	1338

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig	ORF	Start	Stop	match	match	gene name	% sim	% ident	length
ID	ID	(nt)	(nt)	acession					(nt)
259	2	2108	1245	gi 1340128	ORF1	[Staphylococcus aureus]	72	59	864
304	2	285	1094	gi 1205330	glutamine-binding periplasmic protein	[Haemophilus influenzae]	72	52	810
307	10	5326	5039	gi 1070015	protein-dependent	[Bacillus subtilis]	72	53	288
315	1	517	260	gi 1143399	quinol oxidase	[Bacillus subtilis]	72	55	258
316	11	9622	9308	gi 1204445	hypothetical protein (SP:P27857)	[Haemophilus influenzae]	72	52	315
337	3	926	1609	gi 1487433	citrate synthase II	[Bacillus subtilis]	72	55	684
364	7	112538	10493	gi 1510643	ferrous iron transport protein B	[Methanococcus jannaschii]	72	53	2046
409	2	340	1263	gi 1402944	orfRM1 gene product	[Bacillus subtilis]	72	49	924
441	3	2177	1590	gi 1312379	highly conserved among eubacteria [Clostridium acetobutylicum]		72	48	588
					pir S34312 S34312 hypothetical protein V - Clostridium cetobutylicum				
453	6	2654	2505	pir S00601 BXSA	antibacterial protein 3 - Staphylococcus haemolyticus		72	70	150
460	1	2	625	gi 1016162	ABC transporter subunit [Cyanophora paradoxa]		72	51	624
463	1	3253	1628	gi 1666014	The polymorphysm (RFLP) of this gene is associated with usceptibility to essential hypertension. The SA gene product has light homology to acetyl-CoA synthetase [Homo sapiens]		72	60	1626
480	4	3047	3466	gi 1433992	ATP synthase subunit epsilon	[Bacillus subtilis]	72	53	420
502	1	1086	586	gi 1310859	ORF2 [Synechococcus sp.]		72	50	501
519	1	81	1184	gi 11303704	YrKE	[Bacillus subtilis]	72	54	1104
559	1	3	746	gi 1107530	ceuD gene product	[Campylobacter coli]	72	56	744
575	1	1142	573	gi 11303866	YggS	[Bacillus subtilis]	72	56	570
671	1	2	592	gi 1204497	protein-export membrane protein	[Haemophilus influenzae]	72	44	591
679	2	295	1251	gi 1563258	virulence-associated protein E	[Dichelobacter nodosus]	72	52	957
687	2	295	957	gi 11146214	44% identical amino acids with the Escherichia coli smba supress; putative		72	49	663
					[Bacillus subtilis]				
837	1	1	435	gi 11146183	putative	[Bacillus subtilis]	72	54	435
868	1	150	788	gi 1377842	unknown	[Bacillus subtilis]	72	55	639
922	1	130	432	gi 1088269	unknown protein	[Azotobacter vinelandii]	72	58	303
941	1	2	238	gi 1153929	NADPH-sulfite reductase	[Salmonella yphimurium]	72	49	237
980	1	840	421	gi 1853767	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	[Bacillus ubtilis]	72	59	420

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1209	1	383	213	gi 1144735	neurotoxin type B [Clostridium botulinum]	72	44	171
1469	2	671	474	gi 1205458	hypothetical protein (GB:D26562.47) [Haemophilus influenzae]	72	63	198
1956	1	727	365	gi 1154409	hexosephosphate transport protein [Salmonella typhimurium]	72	44	363
				pir B41853 B41853 hexose phosphate transport system regulatory rotein uhpB				
				- Salmonella typhimurium				
2101	1	3	401	gi 1303950	YqiY [Bacillus subtilis]	72	50	399
2503	1	569	399	gi 1149713	formate dehydrogenase [Methanobacterium formicicum] pir A42712 A42712 formate dehydrogenase (EC 1.2.1.2) - ethanobacterium formicicum	72	56	171
2967	1	3	155	gi 1212729	YqhJ [Bacillus subtilis]	72	46	153
3004	1	367	185	gi 1665999	hypothetical protein [Bacillus subtilis]	72	55	183
3109	1	278	141	gi 1413968	ipa-44d gene product [Bacillus subtilis]	72	45	138
3171	1	3	287	gi 1515938	glutamate synthase (ferredoxin) [Synechocystis sp.] pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	72	52	285
3771	1	26	367	gi 11408501	homologous to N-acyl-L-amino acid amidohydrolase of Bacillus stearothermophilus [Bacillus subtilis]	72	63	342
3951	1	1	222	gi 1500409	M. jannaschii predicted coding region MJ1519 [Methanococcus jannaschii]	72	38	222
4190	1	721	362	gi 39956	IIGlc [Bacillus subtilis]	72	57	360
4444	1	3	347	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	72	55	345
6	2	931	1200	gi 1537095	ornithine carbamoyltransferase [Escherichia coli]	71	55	270
11	15	11350	10859	gi 1532309	25 kDa protein [Escherichia coli]	71	47	492
19	2	1248	2435	gi 1244574	D-alanine:D-alanine ligase [Enterococcus hirae]	71	52	1188
21	2	898	1488	gi 1149629	anthranilate synthase component 2 [Leptospira biflexa] pir C32840 C32840 anthranilate synthase (EC 4.1.3.27) component II Leptospira biflexa	71	45	591
34	1	1	567	gi 1303983	YqkF [Bacillus subtilis]	- 71	59	567
37	3	3192	2806	gi 1209681	glutamate-rich protein [Bacillus firmus]	71	50	387
38	18	12250	12462	gi 1927645	arginyl endopeptidase [Porphyromonas gingivalis]	71	50	213
39	3	1246	4431	pir S09411 S094	spoIIIE protein - Bacillus subtilis	71	49	3186
53	14	15770	14760	gi 1142611	branched chain alpha-keto acid dehydrogenase E1-alpha [Bacillus ubtilis]	71	58	1011
54	11	113461	12625	gi 1143014	gnt repressor [Bacillus subtilis]	71	46	837

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
57	7	7152	5860	gi 508175	EIIC domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	71	48	1293
57	18	13897	14334	gi 1063247	high homology to flavohemoprotein (Haemoglobin-like protein) of Alcaligenes eutrophus and Saccharomyces cerevisiae [Bacillus subtilis]	71	56	438
62	16	9831	10955	gi 1303926	YqiG [Bacillus subtilis]	71	54	1125
70	12	8505	8966	gi 147198	phnE protein [Escherichia coli]	71	38	462
86	5	2394	2089	gi 904205	hypothetical protein [Bacillus subtilis]	71	51	306
96	7	7601	8269	gi 709991	hypothetical protein [Bacillus subtilis]	71	49	669
100	6	4822	5931	gi 1060848	Opine dehydrogenase [Arthrobacter sp.]	71	45	1110
103	1	1062	532	gi 143089	iep protein [Bacillus subtilis]	71	41	531
109	18	15312	15695	gi 413985	ipa-61d gene product [Bacillus subtilis]	71	57	384
113	1	630	316	gi 663254	probable protein kinase [Saccharomyces cerevisiae]	71	57	315
114	5	6598	5603	gi 143156	membrane bound protein [Bacillus subtilis]	71	40	996
133	2	3087	1723	gi 1303913	YqhX [Bacillus subtilis]	71	53	1365
149	19	6335	5895	gi 529650	G40P [Bacteriophage SPPI]	71	51	441
154	5	3635	3087	gi 425488	repressor protein [Streptococcus sobrinus]	71	47	549
164	11	11354	11689	gi 49318	ORF4 gene product [Bacillus subtilis]	71	52	336
169	5	1936	2745	gi 1403403	unknown [Mycobacterium tuberculosis]	71	56	810
193	2	272	1234	gi 1303788	YqeH [Bacillus subtilis]	71	49	963
205	1	1743	895	gi 1215694	GlnQ [Mycoplasma pneumoniae]	71	46	849
233	4	1849	2022	gi 633732	ORF1 [Campylobacter jejuni]	71	50	174
237	7	4501	5169	gi 149384	HisIE [Lactococcus lactis]	71	54	669
272	4	2848	2273	gi 709993	hypothetical protein [Bacillus subtilis]	71	48	576
274	2	618	1496	gi 143035	[NAD(P)H:glutaryl-transfer RNA reductase [Bacillus subtilis] pir A35252 A35252 5-aminolevulinate synthase (EC 2.3.1.37) - acillus subtilis]	71	53	879
276	5	3349	2720	gi 303562	ORF210 [Escherichia coli]	71	50	630
287	1	136	660	gi 310634	120 kDa protein [Streptococcus gordonii]	71	53	525
288	6	3322	2771	gi 1256625	putative [Bacillus subtilis]	71	47	552

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
301	6	3492	2461	gi 467417	similar to lysine decarboxylase [Bacillus subtilis]	71	57	1032
306	4	6607	5222	gi 1256618	transport protein [Bacillus subtilis]	71	56	1386
307	2	1536	925	gi 602683	orfC [Mycoplasma capricolum]	71	45	612
310	5	5793	5146	gi 348052	acetoin utilization protein [Bacillus subtilis]	71	51	648
322	1	2	1303	gi 1001819	hypothetical protein [Synecocystis sp.]	71	46	1302
333	4	4171	3995	gi 467473	unknown [Bacillus subtilis]	71	57	177
350	2	548	922	gi 551879	ORF 1 [Lactococcus lactis]	71	55	375
375	4	1860	3071	gi 467447	unknown [Bacillus subtilis]	71	57	1212
380	5	1560	2102	gi 142557	ATP synthase b subunit [Bacillus megaterium]	71	43	543
414	2	251	637	gi 580904	homologous to E.coli rnpA [Bacillus subtilis]	71	49	387
424	1	335	1354	gi 581305	L-lactate dehydrogenase [Lactobacillus plantarum]	71	57	1020
436	4	3701	3270	pir PN0501 PN05	phosphoribosylanthranilate isomerase (EC 5.3.1.24) - Bacillus subtilis (fragment)	71	66	432
482	1	3	1280	gi 410142	ORFX18 [Bacillus subtilis]	71	49	1278
525	3	2272	1844	gi 143370	phosphoribosylpyrophosphate amidotransferase (PUR-F; EC 2.4.2.14) Bacillus subtilis	71	56	429
529	4	2739	2047	gi 606150	ORF_f309 [Escherichia coli]	71	43	693
563	1	22	969	gi 1237015	ORF4 [Bacillus subtilis]	71	53	948
581	1	506	255	gi 1301730	T25G3.2 [Caenorhabditis elegans]	71	47	252
612	2	1068	913	gi 153968	fimbriae 2 [Salmonella typhimurium]	71	55	156
613	1	1	654	gi 466778	lysine specific permease [Escherichia coli]	71	50	654
618	1	1243	623	gi 1146238	poly(A) polymerase [Bacillus subtilis]	71	52	621
630	1	1170	586	gi 1486243	unknown [Bacillus subtilis]	71	53	585
691	1	1126	641	gi 289260	come ORF1 [Bacillus subtilis]	71	51	486
694	2	149	427	gi 12971	NADH dehydrogenase subunit V (AA 1-605) [gallus gallus] ir S10197 S10197	71	47	279
715	2	169	777	gi 1303830	YqfI [Bacillus subtilis]	71	53	609
746	2	1473	970	gi 1377843	unknown [Bacillus subtilis]	71	52	504

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
748	1	1437	802	gi 1405459	YneS [Bacillus subtilis]	71	49	636
753	1	1018	524	gi 1510389	[M. jannaschii predicted coding region MJ0296 [Methanococcus jannaschii]	71	53	495
761	1	3	215	gi 475972	pentafunctional enzyme [Pneumocystis carinii]	71	47	213
783	1	1203	703	gi 536655	ORF YBR244w [Saccharomyces cerevisiae]	71	52	501
800	3	1292	987	gi 1204326	tRNA delta(2)-isopentenylpyrophosphate transferase [Haemophilus influenzae]	71	48	306
806	1	116	286	gi 1419075	cbiM gene product [Methanobacterium thermoautotrophicum]	71	50	171
931	1	973	488	gi 893358	PgsA [Bacillus subtilis]	71	56	486
1041	1	2	262	gi 1408507	pyrimidine nucleoside transport protein [Bacillus subtilis]	71	45	261
1070	1	2	172	gi 709993	hypothetical protein [Bacillus subtilis]	71	46	171
1176	1	57	365	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir[A44756[A44756 Pseudomonas sp. hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88)]	71	49	309
1181	1	366	184	gi 46971	lepIP gene product [Staphylococcus epidermidis]	71	50	183
1281	1	3	290	gi 153016	ORF 419 protein [Staphylococcus aureus]	71	50	288
1348	1	456	229	gi 602683	orfC [Mycoplasma capricolum]	71	48	228
2002	1	756	379	gi 1008177	ORF YJL046w [Saccharomyces cerevisiae]	71	48	378
2119	1	2	217	gi 1046088	arginyl-tRNA synthetase [Mycoplasma genitalium]	71	50	216
2418	1	3	320	gi 1499771	[M. jannaschii predicted coding region MJ0936 [Methanococcus jannaschii]	71	57	318
2961	1	2	187	gi 312443	[carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus aldolyticus]	71	57	186
2999	2	67	306	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	71	43	240
3033	1	2	184	gi 1262335	YmaA [Bacillus subtilis]	71	57	183
3584	1	3	338	gi 401716	beta-isopropylmalate dehydrogenase [Neurospora crassa]	71	55	336
3715	2	743	399	gi 563952	gluconate permease [Bacillus licheniformis]	71	59	345
3785	1	770	387	gi 47382	acyl-CoA-dehydrogenase [Streptomyces purpurascens]	71	57	384
3875	1	541	272	gi 1001541	hypothetical protein [Synecocystis sp.]	71	38	270
4135	1	637	320	gi 142695	[S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase Bacillus megaterium]	71	52	318
4249	1	63	239	gi 1205363	deoxyribose aldolase [Haemophilus influenzae]	71	63	177
4508	1	530	267	gi 1197667	vitellogenin [Anolis pulchellus]	71	46	264

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
6	3	1237	2721	gi 1321788	arginine ornithine antiporter [Clostridium perfringens]	70	54	1485
11	11	6572	7486	gi 216854	P47K [Pseudomonas chlororaphis]	70	41	915
12	1	2890	1481	gi 467330	replicative DNA helicase [Bacillus subtilis]	70	49	1410
15	1	1756	893	gi 451216	Mannosephosphate Isomerase [Streptococcus mutans]	70	46	864
15	2	1277	1050	gi 476092	unknown [Bacillus subtilis]	70	50	228
17	2	2132	1350	gi 145402	choline dehydrogenase [Escherichia coli]	70	52	783
21	1	2	925	gi 149516	anthranilate synthase alpha subunit [Lactococcus lactis] pir S35124 S35124	70	50	924
					anthranilate synthase (EC 4.1.3.27) alpha chain - actococcus lactis subsp. lactis			
25	7	5580	6251	gi 1389549	ORF3 [Bacillus subtilis]	70	52	672
33	6	6071	7423	gi 1303875	YqhB [Bacillus subtilis]	70	51	1353
36	2	959	1594	gi 500755	methyl purine glycosylase [Mus musculus]	70	47	636
38	8	4901	5860	gi 1408507	pyrimidine nucleoside transport protein [Bacillus subtilis]	70	44	960
44	8	5312	5989	gi 1006620	hypothetical protein [Synechocystis sp.]	70	49	678
46	10	8950	10020	gi 1403126	czcD gene product [Alcaligenes eutrophus]	70	45	1071
52	2	2727	1900	gi 1486247	unknown [Bacillus subtilis]	70	53	828
52	6	4048	4656	gi 244501	esterase II=carboxylesterase (EC 3.1.1.1) [Pseudomonas fluorescens, eptide, 218 aa]	70	50	609
56	8	8460	9962	gi 1339951	small subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	70	51	1503
62	1	48	290	gi 142702	A competence protein 2 [Bacillus subtilis]	70	47	243
64	1	1080	541	gi 1204377	molybdopterin biosynthesis protein [Haemophilus influenzae]	70	47	540
70	5	5139	3595	gi 1204834	2',3'-cyclic-nucleotide 2'-phosphodiesterase [Haemophilus influenzae]	70	47	1545
91	4	7793	5466	gi 886471	methionine synthase [Catharanthus roseus]	70	56	2328
96	5	8754	7255	pir B39096 B390	alkaline phosphatase (EC 3.1.3.1) III precursor - Bacillus subtilis	70	54	1500
110	2	767	1300	gi 145294	adenine phosphoribosyl-transferase [Escherichia coli]	70	51	534
116	6	7026	7976	gi 143607	sporulation protein [Bacillus subtilis]	70	50	951
121	8	6401	6988	gi 1107528	ttg start [Campylobacter coli]	70	45	588
131	8	6842	7936	gi 1150454	prolidase PepQ [Lactobacillus delbrueckii]	70	48	1095

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
135	1	2	1489	gi 311309	putative membrane-bound protein with four times repetition of ro-Ser-Ala at the N-terminus; function unknown [Alcaligenes utrophus]	70	49	1488
138	3	418	714	gi 904181	hypothetical protein [Bacillus subtilis]	70	46	297
164	8	9344	9874	gi 49315	ORF1 gene product [Bacillus subtilis]	70	47	531
164	16	15626	16618	gi 1205212	hypothetical protein (GB:D10483_18) [Haemophilus influenzae]	70	50	993
205	2	2735	1803	gi 1215695	peptide transport system protein SapF homolog; SapF homolog [Mycoplasma pneumoniae]	70	47	933
209	3	910	1386	gi 1204665	hypothetical protein (GB:X73124_26) [Haemophilus influenzae]	70	48	477
246	3	340	756	gi 215098	excisionase [Bacteriophage l54a]	70	46	417
263	7	7876	6749	gi 142540	aspartokinase II [Bacillus sp.]	70	51	1128
268	3	3212	4117	gi 1340128	ORF1 [Staphylococcus aureus]	70	50	906
302	6	3201	3827	gi 147782	ruvA protein (gtg start) [Escherichia coli]	70	46	627
302	10	5879	7051	pir C38530 C385	queuine tRNA-ribosyltransferase (EC 2.4.2.29) - Escherichia coli	70	55	1173
313	1	2520	1414	gi 1205934	aminopeptidase a/i [Haemophilus influenzae]	70	46	1107
355	2	379	669	gi 1070013	protein-dependent [Bacillus subtilis]	70	48	291
403	1	1255	629	gi 733147	GumF [Xanthomonas campestris]	70	33	627
444	10	8770	9273	gi 1204752	high affinity ribose transport protein [Haemophilus influenzae]	70	52	504
449	1	2	1243	gi 1619724	MgtE [Bacillus firmus]	70	44	1242
472	1	637	320	gi 727145	open reading frame; putative [Bacillus amyloliquefaciens] pir B29091 B29091	70	41	318
480	2	727	1608	gi 142560	ATP synthase gamma subunit [Bacillus megaterium]	70	44	882
524	1	2	307	gi 602292	RCH2 protein [Brassica napus]	70	45	306
525	1	823	413	gi 143372	phosphoribosyl glycinamide formyltransferase (PUR-N) [Bacillus ubtilis]	70	52	411
565	4	3625	2552	gi 881434	OREP [Bacillus subtilis]	70	51	1074
607	4	829	1284	gi 1511524	hypothetical protein (SP:P37002) [Methanococcus jannaschii]	70	50	456
633	1	1383	703	gi 431231	uracil permease [Bacillus caldolyticus]	70	53	681
646	3	1683	1309	gi 467340	unknown [Bacillus subtilis]	70	49	375
663	1	830	417	gi 1303873	YggZ [Bacillus subtilis]	70	40	414

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
681	1	1488	781	gi 1001678	hypothetical protein [Synechocystis sp.]	70	53	708
708	1	2	448	sp P33940 YOJH_HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION.		70	51	447
725	1	51	722	gi 1001644	hypothetical protein [Synechocystis sp.]	70	48	672
776	1	1371	787	gi 145165	putative [Escherichia coli]	70	47	585
834	1	250	783	gi 552971	NADH dehydrogenase (ndhF) [Vicia faba]	70	47	534
865	2	1585	1379	gi 1204636	ATP-dependent helicase [Haemophilus influenzae]	70	45	207
894	1	535	269	gi 467364	DNA binding protein (probale) [Bacillus subtilis]	70	41	267
919	1	3	317	gi 1314847	CinA [Bacillus subtilis]	70	40	315
944	1	3	572	gi 709991	hypothetical protein [Bacillus subtilis]	70	44	570
988	2	772	605	gi 142441	ORF 3; putative [Bacillus subtilis]	70	50	168
1055	1	3	335	gi 529755	speC [Streptococcus pyogenes]	70	37	333
1093	1	2	904	gi 853754	ABC transporter [Bacillus subtilis]	70	49	903
1109	1	2	310	gi 1001827	hypothetical protein [Synechocystis sp.]	70	42	309
1220	1	468	235	pir S23416 S234	epiB protein - Staphylococcus epidermidis	70	40	234
1279	1	73	348	gi 153015	FenA protein [Staphylococcus aureus]	70	47	276
1336	1	195	542	sp P31776 PBPA_PENICILLIN-BINDING PROTEIN 1A (PBP-1A) (PENICILLIN-BINDING PROTEIN A).		70	50	348
1537	2	232	402	gi 1146181	putative [Bacillus subtilis]	70	50	171
1574	1	451	272	gi 219630	endothelin-A receptor [Homo sapiens]	70	47	180
1640	1	690	346	gi 1146243	22.4% identity with Escherichia coli DNA-damage inducible protein ...; putative [Bacillus subtilis]	70	46	345
2504	1	2	286	gi 495179	transmembrane protein [Lactococcus lactis]	70	51	285
3061	1	564	301	gi 508175	EIRC domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	70	44	264
3128	1	2	199	gi 1340096	unknown [Mycobacterium tuberculosis]	70	51	198
3218	1	3	488	gi 515938	glutamate synthase (ferredoxin) [Synechocystis sp.] pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	70	50	486
3323	1	794	399	gi 1154891	ATP binding protein [Phormidium laminosum]	70	52	396
3679	1	599	399	gi 529385	chromosome condensation protein [Caenorhabditis elegans]	70	30	201

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3841	1	706	398	gi 1208965	hypothetical 23.3 kd protein [Escherichia coli]	70	47	309
3929	1	3	401	gi 149435	putative [Lactococcus lactis]	70	49	399
4044	1	595	374	gi 602031	[similar to trimethylamine DH [Mycoplasma capricolum] pir S49950 S49950 probable trimethylamine dehydrogenase (EC .5.99.7) - Mycoplasma capricolum] (SGC3) (fragment)	70	40	222
4329	1	558	280	gi 1339951	small subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	70	49	279
4422	1	576	289	gi 296464	ATPase [Lactococcus lactis]	70	57	288
4647	1	361	200	gi 166412	NADH-glutamate synthase [Medicago sativa]	70	59	162
16	8	7571	9031	gi 1499620	[M. jannaschii predicted coding region MJ0798 [Methanococcus jannaschii]	69	44	1461
16	9	9080	10033	gi 1353197	[thioredoxin reductase [Eubacterium acidaminophilum]	69	54	954
30	1	1452	727	gi 1204910	hypothetical protein (GB:U14003_302) [Haemophilus influenzae]	69	52	726
38	4	1023	1298	gi 407773	[devA gene product [Anabaena sp.]	69	41	276
44	9	5987	6595	gi 1205920	[molybdate uptake system hydrophilic membrane-bound protein [Haemophilus influenzae]	69	45	609
62	15	9104	9475	gi 385178	unknown [Bacillus subtilis]	69	44	372
66	4	2402	2803	gi 1303893	YqhL [Bacillus subtilis]	69	51	402
67	15	14124	13627	gi 149647	OREZ [Listeria monocytogenes]	69	37	498
67	17	14053	14382	gi 305002	ORE_f356 [Escherichia coli]	69	49	330
67	19	15130	15807	gi 1109684	Prov [Bacillus subtilis]	69	45	678
78	3	1447	2124	gi 1256633	[putative [Bacillus subtilis]	69	53	678
78	4	4513	3725	gi 1303958	YqjG [Bacillus subtilis]	69	32	789
85	4	4521	4213	pir E29326 E293	hypothetical protein (pur operon) - Bacillus subtilis	69	32	309
86	6	3253	2654	gi 1973332	Orfc [Bacillus subtilis]	69	50	600
95	1	96	710	gi 786468	[A11 antigen, sperm tail membrane antigen-putative sucrose-specific hosphotransferase enzyme II homolog [mice, testis, Peptide Partial, 72 aa]	69	43	615
100	7	6023	7426	gi 1205355	[Na+/H+ antiporter [Haemophilus influenzae]	69	39	1404
102	2	2678	1650	gi 561690	[sialoglycoprotease [Pasteurella haemolytica]	69	47	1029
103	8	12241	8537	gi 1009366	[Respiratory nitrate reductase [Bacillus subtilis]	69	54	3705
103	11	14987	12552	gi 710020	[nitrite reductase (nirB) [Bacillus subtilis]	69	51	2436

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
112	111	8708	10168	gi 154411	hexosephosphate transport protein [Salmonella typhimurium]	69	51	1461
				pir D41853 D41853	hexose phosphate transport system protein uhpf - almonella typhimurium			
112	116	16644	17414	gi 1204435	pyruvate formate-lyase activating enzyme [Haemophilus influenzae]	69	50	771
113	2	33	953	gi 290509	o307 [Escherichia coli]	69	43	921
114	2	1537	1058	pir A42771 A427	reticulocyte-binding protein 1 - Plasmodium vivax	69	39	480
121	6	4309	5310	gi 1154633	NrdF [Bacillus subtilis]	69	53	1002
125	2	267	854	gi 413931	lipa-7d gene product [Bacillus subtilis]	69	43	588
149	27	10666	10400	pir S28089 S280	hypothetical protein A - yeast (Zygosaccharomyces bisporus) plasmid pSB3	69	39	267
161	1	1598	813	gi 1205538	hypothetical protein (GB:U14003_302) [Haemophilus influenzae]	69	47	786
165	4	2222	4633	gi 40054	phenylalanyl-tRNA synthetase beta subunit (AA 1-804) [Bacillus btillis]	69	52	2412
169	3	1210	1761	gi 296031	elongation factor Ts [Spirulina platensis]	69	45	552
175	12	8686	8339	gi 732682	FimE protein [Escherichia coli]	69	69	348
190	2	484	1671	spi P17731 HIS8_	HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-PHOSPHATE TRANSAMINASE)	69	48	1188
206	4	6038	5796	gi 1256135	YbbF [Bacillus subtilis]	69	48	243
249	1	636	319	gi 1405456	YneP [Bacillus subtilis]	69	50	318
302	8	4820	5776	gi 1001768	hypothetical protein [Synecocystis sp.]	69	48	957
324	2	7384	3893	gi 1256798	pyruvate carboxylase [Rhizobium etli]	69	53	3492
351	3	2098	1808	gi 1491664	T04H1.4 [Caenorhabditis elegans]	69	30	291
369	3	2075	2305	gi 336458	ORF [Balaenoptera acutorostrata]	69	61	231
392	3	1999	2424	gi 556015	ORF1 [Bacillus subtilis]	69	45	426
410	1	87	779	gi 155611	phosphoglyceromutase [Zymomonas mobilis]	69	58	693
421	1	2085	1129	gi 1276985	arginase [Bacillus caldovelox]	69	54	957
444	8	6713	7741	gi 1221782	purine synthesis repressor [Haemophilus influenzae]	69	40	1029
453	1	828	415	gi 1122758	unknown [Bacillus subtilis]	69	57	414
469	2	3286	2246	gi 1458228	mutY homolog [Homo sapiens]	69	44	1041

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
509	3	1730	1371	gi 49224	URF 4 [Synecococcus sp.]	69	39	360
520	5	3023	2823	gi 726427	similar to D. melanogaster MST101-2 protein (PIR:S34154) Caenorhabditis elegans]	69	39	201
531	1	26	760	gi 509672	repressor protein [Bacteriophage Tuc2009]	69	33	735
589	1	107	253	gi 169101	17.9 kDa heat shock protein (hspl7.9) [Pisum sativum]	69	52	147
594	2	597	1391	gi 142783	DNA photolyase [Bacillus firmus]	69	48	795
604	4	2476	2114	gi 413930	ipa-6d gene product [Bacillus subtilis]	69	45	363
607	1	2	313	gi 1236103	W08D2.3 [Caenorhabditis elegans]	69	47	312
607	2	590	312	gi 536715	ORF YBR275c [Saccharomyces cerevisiae]	69	39	279
734	1	864	433	gi 467327	unknown [Bacillus subtilis]	69	44	432
759	1	3	338	gi 1009367	Respiratory nitrate reductase [Bacillus subtilis]	69	50	336
761	2	392	586	gi 3508	Leucyl-tRNA synthetase (cytoplasmic) [Saccharomyces cerevisiae] 370340	69	46	195
802	1	72	1013	gi 143044	ferrochelatase [Bacillus subtilis]	69	55	942
816	1	2573	1368	gi 1510268	restriction modification system S subunit [Methanococcus jannaschii]	69	45	1206
838	2	133	387	gi 1255371	coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase [Caenorhabditis elegans]	69	46	255
851	2	745	1005	gi 288998	secA gene product [Antithamnion sp.]	69	39	261
867	1	535	269	gi 1070014	protein-dependent [Bacillus subtilis]	69	47	267
995	1	954	478	gi 1205569	transcription elongation factor [Haemophilus influenzae]	69	53	477
999	1	1009	506	gi 899254	predicted trithorax protein [Drosophila virilis]	69	21	504
1127	1	1315	659	gi 1205434	H. influenzae predicted coding region H1191 [Haemophilus influenzae]	69	56	657
1138	1	248	460	gi 1510646	M. jannaschii predicted coding region MJ0568 [Methanococcus jannaschii]	69	48	213
2928	1	3	401	gi 290503	glutamate permease [Escherichia coli]	69	41	399
3090	1	444	223	gi 1204987	DNA polymerase III, alpha chain [Haemophilus influenzae]	69	36	222
3817	1	2	400	gi 1483199	peptide-synthetase [Amycolatopsis mediterranei]	69	45	399
3833	1	667	335	gi 1524193	unknown [Mycobacterium tuberculosis]	69	46	333

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4079	1	747	400	gi 546918	orfY 3' of conK [Bacillus subtilis, E26, Peptide Partial, 140 aa]	69	64	348
				pir S43612 S43612	hypothetical protein Y - Bacillus subtilis			
				sp P40398 YHXD_BACSU	HYPOTHETICAL PROTEIN IN CONK 3'REGION (ORFY FRAGMENT).			
4115	2	215	400	gi 517205	167 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	69	59	186
4139	1	1	333	gi 1208451	hypothetical protein [Synechocystis sp.]	69	36	333
4258	1	457	230	gi 496158	restriction-modification enzyme subunit M1 [Mycoplasma pulmonis]	69	43	228
				pir S49395 S49395	HsdM1 protein - Mycoplasma pulmonis (SGC3)			
4317	1	90	374	gi 413967	ipa-43d gene product [Bacillus subtilis]	69	44	285
4465	1	3	293	gi 396296	similar to phosphotransferase system enzyme II [Escherichia coli]	69	49	291
				sp P32672 PTWC_ECOLI	PTS SYSTEM, FRUCTOSE-LIKE-2 IIC COMPONENT			
					PHOSPHOTRANSFERASE ENZYME II, C COMPONENT).			
3	1	2302	1193	gi 1109685	Prow [Bacillus subtilis]	68	46	1110
15	4	2592	2074	gi 807973	unknown [Saccharomyces cerevisiae]	68	45	519
31	8	6328	8772	gi 290642	ATPase [Enterococcus hirae]	68	48	2445
40	2	1115	750	gi 606342	ORF_o622; reading frame open far upstream of start; possible rameshift, linking to previous ORF [Escherichia coli]	68	55	366
46	9	6886	8415	gi 155276	aldehyde dehydrogenase [Vibrio cholerae]	68	44	1530
48	3	3643	3404	gi 285608	241k polyprotein [Apple stem grooving virus]	68	47	240
48	4	3536	4132	gi 1045937	M. genitalium predicted coding region MG246 [Mycoplasma genitalium]	68	39	597
53	10	11671	10685	gi 1303952	YqjA [Bacillus subtilis]	68	46	987
70	9	7346	8155	gi 147198	phnE protein [Escherichia coli]	68	40	810
89	4	1899	2966	gi 145173	35 kDa protein [Escherichia coli]	68	43	1068
108	1	2187	1150	gi 38722	precursor (aa -20 to 381) [Acinetobacter calcoaceticus] ir A29277 A29277	68	57	1038
					aldose 1-epimerase (EC 5.1.3.3) - Acinetobacter lcoaceticus			
112	5	2666	3622	gi 153724	MalC [Streptococcus pneumoniae]	68	55	957
116	7	7865	8638	gi 143608	sporulation protein [Bacillus subtilis]	68	48	774
118	3	2484	3698	gi 1303805	YqeR [Bacillus subtilis]	68	46	1215
120	2	1424	1594	sp P38038 CYSJ_	SULFITE REDUCTASE (NADPH) FLAVOPROTEIN ALPHA-COMPONENT (EC 1.8.1.2) (SIR-FF).	68	45	171
129	1	1	1011	gi 396307	argininosuccinate lyase [Escherichia coli]	68	50	1011

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
132	3	1867	2739	gi 216267	ORF2 [Bacillus megaterium]	68	48	873
134	2	848	1012	gi 147545	DNA recombinase [Escherichia coli]	68	50	165
141	2	372	614	gi 872116	sti (stress inducible protein) [Glycine max]	68	36	243
149	7	2454	2260	gi 145774	hsp70 protein (dnaK gene) [Escherichia coli]	68	48	195
155	2	1776	1534	gi 216583	ORF1 [Escherichia coli]	68	36	243
158	3	1826	3289	sp P33940 YOJH_	HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION.	68	51	1464
169	6	2749	3318	gi 1403402	unknown [Mycobacterium tuberculosis]	68	46	570
175	10	9158	7365	gi 1072395	phaA gene product [Rhizobium meliloti]	68	51	1794
188	7	4184	5434	gi 11173843	3-ketoacyl-ACP synthase II [Vibrio harveyi]	68	48	1251
189	3	907	1665	gi 467383	DNA binding protein (probable) [Bacillus subtilis]	68	55	759
206	5	7683	6709	gi 1256138	YbbI [Bacillus subtilis]	68	48	975
206	8	110425	12176	gi 452687	pyruvate decarboxylase [Saccharomyces cerevisiae]	68	48	1752
212	8	3421	3648	gi 1369941	cl gene product [Bacteriophage B1]	68	39	228
214	8	5457	6482	gi 1420467	ORF YOR196c [Saccharomyces cerevisiae]	68	45	1026
237	4	2507	3088	gi 149381	Hish [Lactococcus lactis]	68	46	582
243	5	5540	4542	gi 1235684	mevalonate pyrophosphate decarboxylase [Saccharomyces cerevisiae]	68	47	999
262	1	3	164	gi 150974	4-oxalocrotonate tautomerase [Pseudomonas putida]	68	42	162
262	2	1984	1118	gi 1147744	PSR [Enterococcus hirae]	68	49	867
276	6	3702	3139	sp P30750 ABC_E	ATP-BINDING PROTEIN ABC (FRAGMENT).	68	50	564
306	6	6345	5725	gi 1256617	adenine phosphoribosyltransferase [Bacillus subtilis]	68	53	621
333	3	4599	3850	gi 467473	unknown [Bacillus subtilis]	68	45	750
365	6	5017	4838	gi 1130643	T22B3.3 [Caenorhabditis elegans]	68	45	180
376	2	549	1646	gi 1277026	DAPA aminotransferase [Bacillus subtilis]	68	51	1098
405	1	1741	872	gi 1303917	YqiB [Bacillus subtilis]	68	47	870
406	2	853	539	gi 1511513	ABC transporter, probable ATP-binding subunit [Methanococcus jannaschii]	68	44	315
426	6	3558	3391	gi 624632	GltL [Escherichia coli]	68	48	168
438	1	108	329	gi 146923	nitrogenase reductase [Escherichia coli]	68	43	222

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
443	1	476	240	gi 535810	hippuricase [Campylobacter jejuni]	68	42	237
443	2	518	1015	gi 1204742	H. influenzae predicted coding region HI0491 [Haemophilus influenzae]	68	48	498
443	5	4447	3779	gi 809660	deoxyribose-phosphate aldolase [Bacillus subtilis] pir S49455 S49455 deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis	68	55	669
476	2	240	1184	gi 971345	unknown, similar to E.coli cardiolipin synthase [Bacillus subtilis] sp P45860 YWIE_BACSU HYPOTHETICAL 58.2 PROTEIN IN NARI-ACDA NTERGENIC REGION.	68	45	945
486	2	1876	1046	gi 147328	transport protein [Escherichia coli]	68	41	831
517	3	1764	2084	gi 1523809	orf2 [Bacteriophage A2]	68	64	321
572	1	2	571	sp P39237 Y05L_	HYPOTHETICAL 6.8 KD PROTEIN IN NRDC-TK INTERGENIC REGION.	68	47	570
646	1	914	459	gi 413982	ipa-58r gene product [Bacillus subtilis]	68	52	456
659	3	1668	1901	gi 1107541	C33D9.8 [Caenorhabditis elegans]	68	36	234
864	5	1510	1716	gi 145774	hsp70 protein (dnaK gene) [Escherichia coli]	68	48	207
920	1	860	432	gi 1510416	hypothetical protein (SP:P31466) [Methanococcus jannaschii]	68	54	429
952	1	1096	611	gi 603456	reductase [Leishmania major]	68	46	486
970	1	91	402	gi 1354775	pfos/R [Treponema pallidum]	68	46	312
1028	1	1064	534	gi 410117	diaminopimelate decarboxylase [Bacillus subtilis]	68	47	531
1029	1	428	216	gi 1335714	Plasmodium falciparum mRNA for asparagine-rich antigen (clone 17C1)	68	31	213
1058	1	692	348	gi 581649	lepC gene product [Staphylococcus epidermidis]	68	46	345
1096	2	665	465	gi 143434	Rho Factor [Bacillus subtilis]	68	43	201
1308	1	2	694	gi 1469939	group B oligopeptidase PepB [Streptococcus agalactiae]	68	50	693
1679	1	2	238	gi 517205	167 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	68	53	237
2039	1	3	383	gi 153898	transport protein [Salmonella typhimurium]	68	51	381
2077	1	3	326	pir C33496 C334	hisc homolog - Bacillus subtilis	68	47	324
2112	1	613	374	gi 64884	lamin LII [Xenopus laevis]	68	50	240
2273	1	793	398	gi 581648	epiB gene product [staphylococcus epidermidis]	68	45	396
2948	1	2	385	gi 216869	branched-chain amino acid transport carrier [Pseudomonas aeruginosa] pir A38534 A38534 branched-chain amino acid transport protein braz	68	41	384
					Pseudomonas aeruginosa			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
2955	1	768	400	gi1904179	hypothetical protein [Bacillus subtilis]	68	49	369
2981	1	572	288	gi1508979	GTP-binding protein [Bacillus subtilis]	68	48	285
3014	1	584	294	gi11524394	ORF-2 upstream of gbsAB operon [Bacillus subtilis]	68	45	291
3082	1	336	169	gi11204696	fructose-permease IIBC component [Haemophilus influenzae]	68	53	168
3108	1	103	258	gi1217855	heat-shock protein [Arabidopsis thaliana]	68	48	156
3639	1	919	461	gi11510490	nitrate transport permease protein [Methanococcus jannaschii]	68	47	459
3657	1	1	330	gi1155369	PTS enzyme-II fructose [Xanthomonas campestris]	68	48	330
3823	1	780	391	gi1603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis]	68	54	390
				gi1603768 HutI protein, imidazolone-5-propionate hydrolase Bacillus subtilis]				
3982	1	2	277	gi1149435	putative [Lactococcus lactis]	68	47	276
4051	1	1	342	gi1450688	hsdM gene of Ecoprri gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	68	48	342
4089	1	12	209	gi11353678	heavy-metal transporting P-type ATPase [Proteus mirabilis]	68	47	198
4143	1	47	187	gi1603769	HutU protein, urocanase [Bacillus subtilis]	68	55	141
4148	1	2	352	gi1450688	hsdM gene of Ecoprri gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	68	51	351
4173	1	2	382	gi11041097	Pyruvate Kinase [Bacillus psychrophilus]	68	48	381
4182	1	498	250	gi1413968	ipa-44d gene product [Bacillus subtilis]	68	50	249
4362	2	148	318	gi1450688	hsdM gene of Ecoprri gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	68	44	171
5	11	9493	8300	gi1143727	putative [Bacillus subtilis]	67	46	1194
31	11	10318	9833	gi1216746	D-lactate dehydrogenase [Lactobacillus plantarum]	67	41	486
32	3	1560	3155	gi11098557	renal sodium/dicarboxylate cotransporter [Homo sapiens]	67	46	1596
32	5	4945	4145	gi11510720	prephenate dehydratase [Methanococcus jannaschii]	67	51	801
36	5	5350	4268	gi11146216	45% identity with the product of the ORF6 gene from the Erwinia herbicola carotenoid biosynthesis cluster; putative [Bacillus subtilis]	67	58	1083
44	7	4492	5304	gi11006621	hypothetical protein [Synecocystis sp.]	67	43	813

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
56	7	3943	8481	gi 304131	glutamate synthase large subunit precursor [Azospirillum brasilense] pir B46602 B46602 glutamate synthase (NADPH) (EC 1.4.1.13) alpha hain - Azospirillum brasilense	67	52	4539
56	12	13923	14678	gi 1000453	TrcR [Bacillus subtilis]	67	48	756
62	8	5092	4757	gi 1113949	orf3 [Bacillus, C-125, alkali-sensitive mutant 18224, Peptide Mutant, 112 aa]	67	45	336
62	10	7570	6338	gi 854655	Na/H antiporter system [Bacillus alcalophilus]	67	49	1233
99	3	2119	3321	gi 1204349	hypothetical protein (GB:GB:D90212_3) [Haemophilus influenzae]	67	50	1203
102	9	5695	7176	gi 149432	putative [Lactococcus lactis]	67	51	1482
103	13	14549	14049	gi 1408497	LP9D gene product [Bacillus subtilis]	67	48	501
109	15	14821	13982	gi 413976	ipa-52r gene product [Bacillus subtilis]	67	49	840
109	17	14811	15194	gi 413983	ipa-59d gene product [Bacillus subtilis]	67	29	384
121	4	1713	2153	gi 1262335	YmaA [Bacillus subtilis]	67	54	441
122	1	1	1149	gi 143047	ORFB [Bacillus subtilis]	67	35	1149
124	5	4060	3518	gi 556885	Unknown [Bacillus subtilis]	67	47	543
131	2	4584	3589	gi 1046081	hypothetical protein (GB:D26185_10) [Mycoplasma genitalium]	67	30	996
140	3	2899	2297	gi 146549	kdpC [Escherichia coli]	67	45	603
142	4	5409	4198	gi 1212775	GTP cyclohydrolase II [Bacillus amyloliquefaciens]	67	55	1212
147	5	2913	2374	gi 1303709	YrkJ [Bacillus subtilis]	67	44	540
152	8	6341	6673	gi 1377841	unknown [Bacillus subtilis]	67	48	333
161	4	2720	3763	gi 496319	SphX [Synechococcus sp.]	67	47	1044
163	6	1989	3428	gi 595681	2-oxoglutarate/malate translocator [Spinacia oleracea]	67	47	1440
193	3	1351	1626	gi 1511101	shikimate 5-dehydrogenase [Methanococcus jannaschii]	67	53	276
200	2	917	2179	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	67	48	1263
206	10	12445	12801	sp P37347 YECD_HYPOTHETICAL 21.8 KD PROTEIN IN ASPS 5'REGION.		67	47	357
206	11	13047	14432	gi 732813	branched-chain amino acid carrier [Lactobacillus delbrueckii]	67	46	1386
208	2	1321	809	gi 1033037	100 kDa heat shock protein (Hsp100) [Leishmania major]	67	36	513
238	3	1039	2052	gi 809542	CbrB protein [Erwinia chrysanthemi]	67	42	1014

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
246	2	176	367	gi1215098	excisionase [Bacteriophage 154a]	67	37	192
276	2	2260	1412	gi1303560	ORF271 [Escherichia coli]	67	50	849
297	6	2223	3056	gi1142784	CtaA protein [Bacillus firmus]	67	46	834
307	7	5220	4186	gi11070013	protein-dependent [Bacillus subtilis]	67	43	1035
316	1	36	1028	gi11161061	dioxygenase [Methylobacterium extorquens]	67	52	993
324	3	5650	5030	gi11469784	putative cell division protein ftsW [Enterococcus hirae]	67	49	621
336	1	524	264	gi1173122	urea amidolyase [Saccharomyces cerevisiae]	67	45	261
360	1	108	1394	sp P30053 SYH_S	HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.21) (HISTIDINE--TRNA LIGASE) (HISRS).	67	47	1287
364	3	4890	3592	gi1151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756	67	46	1299
					hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.			
365	3	2940	2113	gi11296823	orf2 gene product [Lactobacillus helveticus]	67	47	828
367	2	325	918	gi11039479	ORFU [Lactococcus lactis]	67	47	594
395	3	666	1271	gi11204516	hypothetical protein (GB:U00014_4) [Haemophilus influenzae]	67	55	606
415	1	1800	901	gi1882579	CG Site No. 29739 [Escherichia coli]	67	46	900
419	1	1799	903	gi1520752	putative [Bacillus subtilis]	67	48	897
474	1	2	796	gi1886906	argininosuccinate synthetase [Streptomyces clavuligerus] pir S57659 S57659	67	49	795
					argininosuccinate synthase (EC 6.3.4.5) - treptomyces clavuligerus			
485	2	1921	2226	gi1143434	Rho Factor [Bacillus subtilis]	67	43	306
596	1	1728	865	gi11303853	YqgF [Bacillus subtilis]	67	47	864
700	1	433	218	gi11204628	hypothetical protein (SP:P21498) [Haemophilus influenzae]	67	47	216
806	2	249	647	gi1677947	AppC [Bacillus subtilis]	67	51	399
828	2	340	900	gi1777761	lrrA [Synechococcus sp.]	67	37	561
833	1	1407	916	gi1142996	regulatory protein [Bacillus subtilis]	67	41	492
856	1	1555	779	gi1780224	ZK970.2 [Caenorhabditis elegans]	67	38	777
888	1	1614	850	gi1437315	TTG start codon [Bacillus licheniformis]	67	40	765
1034	1	1190	597	gi11205113	hypothetical protein (GB:119201_15) [Haemophilus influenzae]	67	45	594
1062	1	636	319	gi11303850	YqgC [Bacillus subtilis]	67	41	318
1067	1	918	460	pir A32950 A329	probable reductase protein - Leishmania major	67	54	459

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1358	1	3	293	gi 1001369	hypothetical protein [Synecocystis sp.]	67	44	291
2181	1	3	302	gi 1510416	hypothetical protein (SP:P31466) [Methanococcus jannaschii]	67	48	300
3000	1	1	507	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	67	56	507
3066	1	464	234	gi 308861	GTG start codon [Lactococcus lactis]	67	46	231
3087	1	454	251	gi 1205366	oligopeptide transport ATP-binding protein [Haemophilus influenzae]	67	44	204
3101	1	2	256	gi 1531541	uroporphyrinogen III methyltransferase [Zea mays]	67	55	255
3598	1	728	393	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756	67	56	336
					hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.			
3765	2	584	366	gi 557489	menD [Bacillus subtilis]	67	45	219
3788	1	658	398	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	67	45	261
3883	1	2	265	gi 704397	cystathionine beta-lyase [Arabidopsis thaliana]	67	46	264
3926	1	2	340	gi 1483199	peptide-synthetase [Amycolatopsis mediterranei]	67	44	339
4417	1	82	396	gi 1205337	ribonucleotide transport ATP-binding protein [Haemophilus influenzae]	67	46	315
2	3	3075	3989	gi 535348	CodV [Bacillus subtilis]	66	42	915
15	6	2273	2542	gi 46491	SmtB [Synecococcus PCC7942]	66	37	270
31	9	8059	7826	gi 292046	mucin [Homo sapiens]	66	44	234
31	10	9034	9258	gi 1204545	mercury scavenger protein [Haemophilus influenzae]	66	48	225
32	6	6347	5253	gi 998342	inducible nitric oxide synthase [Gallus gallus]	66	47	1095
44	13	8856	10124	gi 1510751	molybdenum cofactor biosynthesis moeA protein [Methanococcus jannaschii]	66	46	1269
48	2	1276	2868	gi 150209	ORF 1 [Mycoplasma mycoides]	66	40	1593
58	8	7178	8428	gi 665999	hypothetical protein [Bacillus subtilis]	66	47	1251
62	7	5143	4370	gi 1072398	phaD gene product [Rhizobium meliloti]	66	40	774
70	14	11693	10998	gi 809660	deoxyribose-phosphate aldolase [Bacillus subtilis] pir S49455 S49455	66	55	696
					deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis			
76	1	1	1305	gi 142440	ATP-dependent nuclease [Bacillus subtilis]	66	42	1305
91	6	9236	8205	gi 704397	cystathionine beta-lyase [Arabidopsis thaliana]	66	43	1032
102	5	3810	3265	gi 1204323	hypothetical protein (SP:P31805) [Haemophilus influenzae]	66	41	546

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
103	4	3418	2732	gi 1971344	nitrate reductase gamma subunit [Bacillus subtilis] sp P42177 NARI_BACSU	66	48	687
					NITRATE REDUCTASE GAMMA CHAIN (EC 1.7.99.4). gi 1009369 Respiratory			
					nitrate reductase [Bacillus subtilis] (SUB -160)			
109	6	4243	4674	gi 170886	glucosamine-6-phosphate deaminase [Candida albicans] pir A46652 A46652	66	45	432
					glucosamine-6-phosphate isomerase (EC 5.3.1.10) - east (Candida albicans)			
112	17	17491	17712	gi 1323179	ORF YGR111w [Saccharomyces cerevisiae]	66	33	222
116	2	4667	2637	gi 1491813	gamma-glutamyltranspeptidase [Bacillus subtilis]	66	43	2031
150	5	3189	2989	gi 1146224	putative [Bacillus subtilis]	66	30	201
172	5	3264	3662	gi 1755152	highly hydrophobic integral membrane protein [Bacillus subtilis]	66	41	399
					sp P42953 TAGG_BACSU TEICHOIC ACID TRANSLOCATION PERMEASE PROTEIN AGG.			
174	5	4592	3723	gi 1146241	pantothenate synthetase [Bacillus subtilis]	66	49	870
175	4	3209	2880	gi 1642655	unknown [Rhizobium meliloti]	66	29	330
175	11	8743	7994	gi 1854655	Na/H antiporter system [Bacillus alcalophilus]	66	43	750
190	5	7079	5727	gi 451072	di-tripeptide transporter [Lactococcus lactis]	66	40	1353
195	15	13919	13713	gi 1322411	unknown [Mycobacterium tuberculosis]	66	42	207
217	3	2822	2595	gi 1143542	alternative stop codon [Rattus norvegicus]	66	36	228
233	9	7133	6135	gi 1458327	F08F3.4 gene product [Caenorhabditis elegans]	66	47	999
238	1	43	1041	gi 1809541	CbrA protein [Erwinia chrysanthemi]	66	42	999
241	1	2102	1053	gi 153067	peptidoglycan hydrolase [Staphylococcus aureus]	66	53	1050
261	1	1178	648	gi 1510859	M. jannaschii predicted coding region MJ0790 [Methanococcus jannaschii]	66	40	531
263	3	3731	2973	gi 1205865	tetrahydrodipicolinate N-succinyltransferase [Haemophilus influenzae]	66	47	759
272	8	6548	5484	gi 1882101	high affinity nickel transporter [Alcaligenes eutrophus]	66	44	1065
					sp P23516 HoxN ALCEU HIGH-AFFINITY NICKEL TRANSPORT PROTEIN.			
276	3	2805	2104	gi 1208965	hypothetical 23.3 kd protein [Escherichia coli]	66	47	702
278	2	2830	1784	gi 1488662	phosphatase-associated protein [Bacillus subtilis]	66	48	1047
278	3	3830	2952	gi 303560	ORF271 [Escherichia coli]	66	45	879
279	2	3894	2218	gi 1185289	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1- carboxylate synthase [Bacillus subtilis]	66	48	1677
288	4	2535	2275	gi 1256625	putative [Bacillus subtilis]	66	42	261
292	2	1133	942	gi 1511604	M. jannaschii predicted coding region MJ1651 [Methanococcus jannaschii]	66	30	192

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
294	1	1116	559	gi 216314	esterase [Bacillus stearothermophilus]	66	45	558
297	4	2913	1978	gi 994794	cytochrome a assembly facto [Bacillus subtilis] sp P24009 COXX_BACSU PROBABLE CYTOCHROME C OXIDASE ASSEMBLY FACTOR.	66	45	936
316	4	2053	2682	gi 1107839	alginate lyase [Pseudomonas aeruginosa]	66	40	630
338	4	2460	2302	gi 520750	biotin synthetase [Bacillus sphaericus]	66	58	159
339	1	1214	735	gi 467468	7, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase [Bacillus ubtilis]	66	52	480
363	1	3	863	gi 581649	epiC gene product [Staphylococcus epidermidis]	66	47	861
366	2	232	483	gi 1103505	unknown [Schizosaccharomyces pombe]	66	53	252
367	4	2468	1845	sp P20692 TYRA_	PREPHENATE DEHYDROGENASE (EC 1.3.1.12) (PDH).	66	50	624
372	3	2150	1599	gi 467416	unknown [Bacillus subtilis]	66	38	552
378	1	212	1009	gi 147309	purine nucleoside phosphorylase [Escherichia coli]	66	50	798
401	1	1	462	gi 388263	p-aminobenzoic acid synthase [Streptomyces griseus] pir JN0531 JN0531 p- aminobenzoic acid synthase - Streptomyces riseus	66	46	462
404	7	4826	5254	gi 606744	cytidine deaminase [Bacillus subtilis]	66	51	429
411	2	1738	1103	gi 1460081	unknown [Mycobacterium tuberculosis]	66	44	636
420	1	2	541	gi 1046024	Na+ ATPase subunit J [Mycoplasma genitalium]	66	49	540
431	1	1	858	gi 1500008	M. jannaschii predicted coding region MJ1154 [Methanococcus jannaschii]	66	50	858
443	7	5679	5299	gi 852076	MrgA [Bacillus subtilis]	66	46	381
444	3	3405	2413	gi 153047	lysostaphin (ttg start codon) [Staphylococcus simulans] pir A25881 A25881 lysostaphin precursor - Staphylococcus simulans sp P10547 LSTP_STASI LYSOSTAPHIN PRECURSOR (EC 3.5.1.-).	66	51	993
561	1	956	480	gi 1204905	DNA-3-methyladenine glycosidase I [Haemophilus influenzae]	66	45	477
562	3	1066	1383	gi 1046082	M. genitalium predicted coding region MG372 [Mycoplasma genitalium]	66	52	318
576	1	11	724	gi 305014	ORF_o234 [Escherichia coli]	66	43	714
577	3	1190	903	gi 1001353	hypothetical protein [Synecocystis sp.]	66	52	288
584	1	2	331	sp P24204 YEBA_	HYPOTHETICAL 46.7 KD PROTEIN IN MSBB-RUVB INTERGENIC REGION (ORFU).	66	48	330
592	1	1410	706	gi 928839	ORF266; putative [Lactococcus lactis phage BK5-T]	66	51	705
601	1	1433	720	gi 1488695	novel antigen; orf-2 [Staphylococcus aureus]	66	55	714

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig	ORF	Start	Stop	match	match	gene name	% sim	% ident	length
ID	ID	(nt)	(nt)	acession					(nt)
619	3	468	845	gi 746573		[similar to M. musculus transport system membrane protein, Nkamp PIR:A40739) and S. cerevisiae SMF1 protein (PIR:A45154) Caenorhabditis elegans]	66	45	378
706	2	561	355	gi 1804808		[unknown protein [Rattus norvegicus]	66	46	207
734	2	673	512	gi 1519085		[phosphatidylcholine binding immunoglobulin heavy chain IgM variable region [Mus musculus]	66	60	162
740	1	3	317	gi 1209272		[argininosuccinate lyase [Campylobacter jejuni]	66	42	315
764	1	310	747	gi 435296		[alkaline phosphatase like protein [Lactococcus lactis] pir S39339 S39339 alkaline phosphatase-like protein - Lactococcus actis	66	42	438
852	1	338	171	gi 536955		[CG Site No. 361 [Escherichia coli]	66	43	168
886	1	3	158	gi 289272		[ferrichrome-binding protein [Bacillus subtilis]	66	44	156
889	1	462	232	gi 1833061		[HCMVUL77 (AA 1-642) [Human cytomegalovirus]	66	66	231
893	1	2	247	gi 149008		[putative [Helicobacter pylori]	66	45	246
900	1	1425	733	gi 580842		[F3 [Bacillus subtilis]	66	51	693
906	2	2300	1473	gi 790945		[aryl-alcohol dehydrogenase [Bacillus subtilis]	66	53	828
947	1	79	549	gi 410117		[diaminopimelate decarboxylase [Bacillus subtilis]	66	47	471
950	1	1100	552	gi 48713		[orf145 [Staphylococcus aureus]	66	35	549
955	2	89	475	gi 1204390		[uridine kinase (uridine monophosphokinase) [Haemophilus influenzae]	66	50	387
981	2	1308	997	gi 457146		[rhoptyr protein [Plasmodium yoelii]	66	38	312
986	1	25	315	gi 305002		[ORF_f356 [Escherichia coli]	66	31	291
1057	1	3	203	gi 1303853		[YggF [Bacillus subtilis]	66	40	201
1087	1	1	294	gi 575913		[unknown [Saccharomyces cerevisiae]	66	53	294
1105	1	1	231	gi 1045799		[methylgalactoside permease ATP-binding protein [Mycoplasma genitalium]	66	46	231
1128	1	2	574	gi 1001493		[hypothetical protein [Synechocystis sp.]	66	46	573
1150	1	498	250	gi 1499034		[M. jannaschii predicted coding region MJ0255 [Methanococcus jannaschii]	66	40	249
1180	2	707	453	gi 215908		[DNA polymerase (g43) [Bacteriophage T4]	66	46	255
1208	1	1123	587	gi 1256653		[DNA-binding protein [Bacillus subtilis]	66	58	537
1342	1	1	402	gi 1208474		[hypothetical protein [Synechocystis sp.]	66	53	402
1761	2	589	398	gi 215811		[tail fiber protein [Bacteriophage T3]	66	50	192

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1983	1	499	251	gi 1045935	DNA helicase II [Mycoplasma genitalium]	66	40	249
2103	2	176	400	gi 929798	precursor for the major merozoite surface antigens [Plasmodium alciparum]	66	46	225
2341	1	373	188	gi 1256623	exodeoxyribonuclease [Bacillus subtilis]	66	38	186
2458	1	325	164	gi 1019410	unknown [Schizosaccharomyces pombe]	66	47	162
2505	1	468	235	gi 1510394	putative transcriptional regulator [Methanococcus jannaschii]	66	39	234
2525	1	558	280	gi 1000695	cytotoxin L [Clostridium sordellii]	66	44	279
2935	1	3	275	gi 765073	autolysin [Staphylococcus aureus]	66	47	273
3005	1	114	305	gi 1205784	heterocyst maturation protein [Haemophilus influenzae]	66	46	192
3048	1	80	277	gi 1303813	Yqew [Bacillus subtilis]	66	42	198
3071	1	1	189	gi 1070014	protein-dependent [Bacillus subtilis]	66	41	189
3081	1	404	225	gi 984212	unknown [Schizosaccharomyces pombe]	66	44	180
3090	2	580	386	gi 1204987	DNA polymerase III, alpha chain [Haemophilus influenzae]	66	48	195
3318	1	1	387	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	66	49	387
3739	1	798	400	gi 1109684	ProV [Bacillus subtilis]	66	47	399
3796	1	402	202	gi 853760	acyl-CoA dehydrogenase [Bacillus subtilis]	66	60	201
3924	1	595	347	gi 563952	gluconate permease [Bacillus licheniformis]	66	46	249
4240	1	3	350	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	66	51	348
4604	1	7	234	pir A26713 BHHC	hemocyanin subunit II - Atlantic horseshoe crab	66	46	228
4	9	8845	9750	gi 145646	cynR [Escherichia coli]	65	35	906
6	5	2708	3565	gi 887824	ORF o310 [Escherichia coli]	65	47	858
13	1	1993	998	gi 143402	recombination protein (ttg start codon) [Bacillus subtilis] gi 1303923 RecN [Bacillus subtilis]	65	44	996
15	7	2493	3524	gi 1403126	czcD gene product [Alcaligenes eutrophus]	65	38	1032
18	3	1908	1372	gi 349187	acyltransferase [Saccharomyces cerevisiae]	65	50	537
21	3	1467	2492	gi 149518	phosphoribosyl anthranilate transferase [Lactococcus lactis] pir S35126 S35126 anthranilate phosphoribosyltransferase (EC .4.2.18) - Lactococcus lactis subsp. lactis	65	52	1026
25	4	3374	4312	gi 1502420	malonyl-CoA:Acyl carrier protein transacylase [Bacillus subtilis]	65	44	939

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
27	2	390	626	gi 1212729	YqhJ [Bacillus subtilis]	65	45	237
31	12	11040	10387	gi 509245	D-hydroxyisocaproate dehydrogenase [Lactobacillus delbrueckii]	65	41	654
38	24	19172	19528	gi 547519	H-protein [Flaveria cronquistii]	65	41	357
44	2	790	1746	gi 405882	lyeK [Escherichia coli]	65	46	957
44	12	9356	8832	gi 1205905	Imolybdenum cofactor biosynthesis protein [Haemophilus influenzae]	65	50	525
45	8	6635	7588	gi 493074	ApbA protein [Salmonella typhimurium]	65	46	954
51	2	580	1503	gi 580897	OppB gene product [Bacillus subtilis]	65	45	924
52	1	225	953	gi 1205518	NAD(P)H-flavin oxidoreductase [Haemophilus influenzae]	65	45	729
55	4	1339	1058	pir A44459 A444	troponin T beta TnT-5 - rabbit	65	41	282
67	9	7421	8272	gi 143607	sporulation protein [Bacillus subtilis]	65	42	852
73	5	4446	5375	gi 1204896	lysophospholipase L2 [Haemophilus influenzae]	65	37	930
74	1	954	478	gi 1204844	H. influenzae predicted coding region HI0594 [Haemophilus influenzae]	65	50	477
77	1	2	757	gi 1046082	M. genitalium predicted coding region MG372 [Mycoplasma genitalium]	65	46	756
77	2	795	1433	gi 1222116	permease [Haemophilus influenzae]	65	37	639
81	3	4728	3454	gi 1001708	hypothetical protein [Synechocystis sp.]	65	49	1275
91	7	8548	8357	gi 1399263	cystathionine beta-lyase [Emericella nidulans]	65	40	192
98	3	1608	1988	gi 467423	unknown [Bacillus subtilis]	65	38	381
98	4	2250	2987	gi 467424	unknown [Bacillus subtilis]	65	45	738
102	3	2598	2119	gi 1511532	N-terminal acetyltransferase complex, subunit ARD1 [Methanococcus jannaschii]	65	39	480
102	4	3647	2862	gi 1204637	H. influenzae predicted coding region HI0388 [Haemophilus influenzae]	65	32	786
103	9	10851	9841	gi 142695	S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase Bacillus megaterium	65	47	1011
103	10	10439	10119	gi 710021	nitrite reductase (nirD) [Bacillus subtilis]	65	51	321
106	2	262	1140	gi 39881	ORF 311 (AA 1-311) [Bacillus subtilis]	65	44	879
109	5	3909	4268	gi 1204399	glucosamine-6-phosphate deaminase protein [Haemophilus influenzae]	65	44	360
109	10	7165	8595	gi 536955	CG Site No. 361 [Escherichia coli]	65	41	1431

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
110	4	3688	3915	gi 407881	stringent response-like protein [Streptococcus equisimilis]	65	45	228
				pir S39975 S39975	stringent response-like protein - Streptococcus equisimilis			
110	5	3882	4295	gi 407880	ORF1 [Streptococcus equisimilis]	65	50	414
110	6	4231	4380	gi 11139574	Orf2 [Streptomyces griseus]	65	56	150
112	10	9218	8640	gi 1204571	H. influenzae predicted coding region HI0318 [Haemophilus influenzae]	65	52	579
112	12	12049	11288	gi 710496	transcriptional activator protein [Bacillus brevis]	65	32	762
125	1	2	202	gi 11151158	repeat organellar protein [Plasmodium chabaudi]	65	39	201
126	1	3	422	gi 37589	precursor [Homo sapiens]	65	46	420
127	11	10733	12658	gi 1064809	homologous to sp:HTRA_ECOLI [Bacillus subtilis]	65	41	1926
143	8	7543	7004	gi 1216513	mutator mutT (AT-GC transversion) [Escherichia coli]	65	56	540
145	5	3587	3838	gi 1209768	D02_orf569 [Mycoplasma pneumoniae]	65	27	252
150	4	3482	2841	gi 11146225	putative [Bacillus subtilis]	65	37	642
166	1	3858	1948	gi 1148304	beta-1,4-N-acetylmuramoylhydrolase [Enterococcus hirae] pir A42296 A42296 lysozyme 2 (EC 3.2.1.-) precursor - Enterococcus irae (ATCC 9790)	65	50	1911
188	6	3195	4178	gi 151943	ORF3; putative [Rhodobacter capsulatus]	65	46	984
189	9	4982	4785	gi 58812	ORF IV (AA 1-489) [Figwort mosaic virus]	65	40	198
195	6	7908	5272	gi 1145220	alanyl-tRNA synthetase [Escherichia coli]	65	49	2637
195	7	10599	8104	gi 882711	exonuclease V alpha-subunit [Escherichia coli]	65	38	2496
206	16	16896	18191	gi 408115	ornithine acetyltransferase [Bacillus subtilis]	65	53	1296
217	4	3844	3215	gi 1205974	5'guanylate kinase [Haemophilus influenzae]	65	41	630
220	4	5265	3751	gi 580920	rodD (gtaA) polypeptide (AA 1-673) [Bacillus subtilis] pir S06048 S06048 probable rodD protein - Bacillus subtilis sp P13484 TAGE_BACSU PROBABLE POLY(GLYCEROL-PHOSPHATE) LPHA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS ROTENIN E).	65	40	1515
236	5	2327	3709	gi 1146200	DNA or RNA helicase, DNA-dependent ATPase [Bacillus subtilis]	65	46	1383
237	3	1902	2513	gi 149379	HisBd [Lactococcus lactis]	65	46	612
241	4	4968	4195	gi 1205308	ribonuclease HII (EC 31264) (RNASE HII) [Haemophilus influenzae]	65	50	774
252	1	1278	940	gi 1204989	hypothetical protein (GB:U00022_9) [Haemophilus influenzae]	65	40	339
261	5	4780	3794	gi 145927	fecD [Escherichia coli]	65	43	987

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
274	1	3	278	gi 496558	orfX [Bacillus subtilis]	65	42	276
301	2	982	815	gi 467418	unknown [Bacillus subtilis]	65	45	168
307	4	3586	2864	gi 1070014	protein-dependent [Bacillus subtilis]	65	40	723
335	2	2286	1399	gi 146913	N-acetylglucosamine transport protein [Escherichia coli] pir B29895 WQEC2N (phosphotransferase system enzyme II (EC .7.1.69), N-acetylglucosamine-specific - Escherichia coli sp P09323 PTAA_ECOLI PTS SYSTEM, N-ACETYLGLUCOSAMINE-SPECIFIC IIABC COMPONENT (EIIA	65	50	888
338	5	4120	3170	gi 1277029	biotin synthase [Bacillus subtilis]	65	49	951
343	3	1490	2800	gi 143264	membrane-associated protein [Bacillus subtilis]	65	48	1311
344	4	2761	2531	gi 1050540	tRNA-glutamine synthetase [Lupinus luteus]	65	34	231
358	3	3421	3621	gi 1146220	NAD+ dependent glycerol-3-phosphate dehydrogenase [Bacillus subtilis]	65	47	201
364	1	238	699	gi 1340128	ORF1 [Staphylococcus aureus]	65	51	462
379	1	1	576	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis] pir A27650 A27650 regulatory protein phoR - Bacillus subtilis sp P23545 PHOR_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOR (EC 2.7.3.-).	65	40	576
379	3	3666	4346	gi 143268	dihydroliipoamide transsuccinylase (odhB; EC 2.3.1.61) [Bacillus ubtilis]	65	50	681
428	1	187	483	gi 1420465	ORF YOR195w [Saccharomyces cerevisiae]	65	45	297
438	2	272	838	gi 143498	degS protein [Bacillus subtilis]	65	38	567
444	11	9280	10215	gi 1204756	ribokinase [Haemophilus influenzae]	65	47	936
449	2	1241	1531	gi 599848	Na/H antiporter homolog [Lactococcus lactis]	65	41	291
478	2	1452	865	gi 1045942	glycyl-tRNA synthetase [Mycoplasma genitalium]	65	39	588
479	1	1032	517	gi 1498192	putative [Pseudomonas aeruginosa]	65	40	516
480	6	4312	5637	gi 415662	UDP-N-acetylglucosamine 1-carboxyvinyl transferase [Acinetobacter alcoaceticus]	65	48	1326
484	1	2	430	gi 146551	transmembrane protein (kdpD) [Escherichia coli]	65	44	429
499	1	54	932	gi 603456	reductase [Leishmania major]	65	53	879
505	1	914	459	gi 1518853	OafA [Salmonella typhimurium]	65	39	456
571	2	1509	883	gi 49399	open reading frame upstream glnE [Escherichia coli] ir S37754 S37754 hypothetical protein XE (glnE 5' region) - cherichia coli	65	44	627
611	2	506	270	gi 10961	RAP-2 [Plasmodium falciparum]	65	40	237

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
705	1	564	283	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	65	52	282
712	1	1	177	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	65	37	177
712	2	196	354	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	65	37	159
743	1	2	631	gi 310631	ATP binding protein [Streptococcus gordonii]	65	45	630
749	2	393	779	gi 467374	single strand DNA binding protein [Bacillus subtilis]	65	29	387
762	1	1698	850	gi 160399	multidrug resistance protein [Plasmodium falciparum]	65	48	849
788	1	85	315	gi 1129096	unknown protein [Bacillus sp.]	65	35	231
850	1	1	408	gi 1006604	hypothetical protein [Synecocystis sp.]	65	37	408
908	1	1	444	gi 1199546	2362 [Saccharomyces cerevisiae]	65	46	444
925	1	1	174	gi 1256653	DNA-binding protein [Bacillus subtilis]	65	54	174
1031	1	26	232	gi 238657	AppC=cytochrome d oxidase, subunit I homolog [Escherichia coli, K12, eptide, 514 aa]	65	47	207
1037	1	414	262	gi 1491813	gamma-glutamyltranspeptidase [Bacillus subtilis]	65	46	153
1053	1	348	175	gi 642655	unknown [Rhizobium meliloti]	65	34	174
1149	1	1399	752	gi 1162980	ribulose-5-phosphate 3-epimerase [Spinacia oleracea]	65	48	648
1214	1	881	495	gi 1205959	lactam utilization protein [Haemophilus influenzae]	65	45	387
1276	1	476	276	pir S35493 S354	site-specific DNA-methyltransferase StsI (EC 2.1.1.-) - Streptococcus sanguis	65	35	201
1276	2	900	577	gi 473794	'ORF' [Escherichia coli]	65	34	324
2057	1	272	138	gi 633699	TrsH [Yersinia enterocolitica]	65	21	135
2521	1	336	169	gi 1045789	hypothetical protein (GB:U14003_76) [Mycoplasma genitalium]	65	41	168
2974	1	590	297	gi 152052	enantiomerase-selective amidase [Rhodococcus sp.]	65	45	294
3031	1	306	154	pir JQ1024 JQ10	hypothetical 30K protein (DmrP140 5' region) - fruit fly (Drosophila melanogaster)	65	45	153
3069	1	3	278	gi 144906	product homologous to E.coli thioredoxin reductase: J.Biol.Chem. 1988) 263:9015-9019, and to F52a protein of alkyl hydroperoxide eductase from S.typhimurium: J.Biol.Chem. (1990) 265:10535-10540; pen reading frame A [Clostridium pasteurianum]	65	46	276
3146	1	282	142	gi 49315	ORF1 gene product [Bacillus subtilis]	65	47	141
3170	1	679	341	gi 1507711	indolepyruvate decarboxylase [Erwinia herbicola]	65	44	339

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3546	1	1	303	gi 450688	hsdM gene of Ecopr1 gene product [Escherichia coli] pir S38437 S38437 hsdM	65	42	303
					protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)			
3782	1	2	328	gi 166412	NADH-glutamase synthase [Medicago sativa]	65	42	327
3990	1	374	189	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	65	53	186
4032	1	613	308	gi 1323127	ORF YGR087c [Saccharomyces cerevisiae]	65	50	306
4278	2	726	364	gi 1197667	vitellogenin [Anolis pulchellus]	65	42	363
19	4	4259	5518	gi 145727	dead [Escherichia coli]	64	45	1260
19	6	7639	6926	gi 1016232	ycf27 gene product [Cyanophora paradoxa]	64	36	714
20	8	7053	6454	gi 765073	autolysin [Staphylococcus aureus]	64	47	600
31	13	112706	11537	gi 414009	ipa-85d gene product [Bacillus subtilis]	64	45	1170
33	4	2388	4364	gi 1204696	fructose-permease IIBC component [Haemophilus influenzae]	64	47	1977
36	3	1871	3013	gi 290503	glutamase permease [Escherichia coli]	64	40	1143
37	6	4065	4409	gi 39815	orf 2 gene product [Bacillus subtilis]	64	46	345
45	9	7852	8760	gi 1230585	nucleotide sugar epimerase [Vibrio cholerae O139]	64	53	909
53	3	1540	1899	gi 1303961	YqjJ [Bacillus subtilis]	64	50	360
56	6	4793	3855	gi 457514	glcC [Bacillus subtilis]	64	45	939
56	24	30002	30247	gi 470331	similar to zinc fingers [Caenorhabditis elegans]	64	42	246
62	4	2759	2421	gi 642655	unknown [Rhizobium meliloti]	64	28	339
85	6	7178	6027	gi 457702	5-aminimidazole ribonucleotide-carboxylase [Pichia methanolica]	64	46	1152
					pir S39112 S39112 phosphoribosylaminoimidazole carboxylase (EC .1.1.21) - yeast (Pichia methanolica)			
96	9	9251	10030	gi 1511513	ABC transporter, probable ATP-binding subunit [Methanococcus jannaschii]	64	42	780
100	1	1	600	gi 765073	autolysin [Staphylococcus aureus]	64	44	600
106	5	3868	4854	gi 466778	lysine specific permease [Escherichia coli]	64	46	987
123	2	838	554	gi 467484	unknown [Bacillus subtilis]	64	47	285
127	8	7514	7810	gi 210061	serotype-specific antigen [African horse sickness virus] pir S27891 S27891 capsid protein VP2 - African horse sickness virus	64	28	297
131	7	7134	6721	gi 1511160	M. jannaschii predicted coding region MJ1163 [Methanococcus jannaschii]	64	46	414

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
142	5	5455	4817	gi 1173517	riboflavin synthase alpha subunit [Actinobacillus pleuropneumoniae]	64	44	639
143	1	709	356	pir A32950 A329	probable reductase protein - Leishmania major	64	52	354
149	10	3555	3295	gi 398151	major surface antigen MSG2 [Pneumocystis carinii]	64	44	261
154	4	3134	2307	gi 984587	DinP [Escherichia coli]	64	50	828
161	5	3855	4880	gi 903304	ORF72 [Bacillus subtilis]	64	37	1026
165	1	33	791	gi 467483	unknown [Bacillus subtilis]	64	38	759
175	6	6355	4844	gi 1072398	phad gene product [Rhizobium meliloti]	64	42	1512
188	3	2042	2500	gi 1001961	MHC class II analog [Staphylococcus aureus]	64	45	459
195	14	13667	13446	gi 396380	No definition line found [Escherichia coli]	64	47	222
206	15	16429	16938	gi 304134	argC [Bacillus stearothermophilus]	64	49	510
215	1	560	282	gi 142359	ORF 6 [Azotobacter vinelandii]	64	39	279
243	7	7818	6928	gi 414014	lipa-90d gene product [Bacillus subtilis]	64	49	891
258	2	1330	845	gi 664754	PI7 [Listeria monocytogenes]	64	38	486
259	1	462	232	gi 1499663	M. jannaschii predicted coding region MJ0837 [Methanococcus jannaschii]	64	52	231
263	6	6565	5567	gi 142828	aspartate semialdehyde dehydrogenase [Bacillus subtilis]	64	48	999
				sp Q04797 DHAS_BACSU ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (EC .2.1.11)				
				(ASA DEHYDROGENASE).				
271	1	3	1163	gi 467091	hflX; B2235_C2_202 [Mycobacterium leprae]	64	44	1161
280	1	173	1450	gi 1303839	YqfR [Bacillus subtilis]	64	43	1278
293	1	2532	1267	gi 147345	primosomal protein n' [Escherichia coli]	64	45	1266
295	2	742	1488	gi 459266	Potential membrane spanning protein [Staphylococcus hominis]	64	39	747
				pir S42932 S42932 potential membrane spanning protein - taphylococcus hominis				
301	5	1625	1446	gi 580835	lysine decarboxylase [Bacillus subtilis]	64	35	180
315	4	5064	3949	gi 143396	quinol oxidase [Bacillus subtilis]	64	45	1116
321	1	1264	635	gi 710496	transcriptional activator protein [Bacillus brevis]	64	41	630
333	5	4520	4239	gi 1314295	ORF2; putative 19 kDa protein [Listeria monocytogenes]	64	43	282
342	1	1	549	gi 142940	ftsA [Bacillus subtilis]	64	38	549
353	3	2878	2324	gi 537049	ORF_o470 [Escherichia coli]	64	44	555

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
379	2	827	3658	pir S25295 A328	oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) - Bacillus subtilis	64	47	2832
404	6	4429	4839	pir A36933 A369	diacylglycerol kinase homolog - Streptococcus mutans	64	35	411
407	1	2020	1133	gil 969026	OrfX [Bacillus subtilis]	64	41	888
425	1	1109	591	gil 1146177	phosphotransferase system glucose-specific enzyme II [Bacillus subtilis]	64	44	519
443	6	4082	4798	gil 147309	purine nucleoside phosphorylase [Escherichia coli]	64	51	717
450	2	1035	1604	gil 606376	ORF_o162 [Escherichia coli]	64	38	570
470	5	1680	6107	gil 1369948	host interacting protein [Bacteriophage B1]	64	45	4428
486	4	1911	1471	gil 1205582	spermidine/putrescine transport system permease protein [Haemophilus influenzae]	64	35	441
497	1	2217	1159	spi P36929 FMU_E	FMU PROTEIN.	64	38	1059
501	1	3	410	gil 142450	ahrC protein [Bacillus subtilis]	64	38	408
514	1	3	290	gil 1204496	H. influenzae predicted coding region HI0238 [Haemophilus influenzae]	64	34	288
551	4	3162	3323	gil 1204511	bacterioferritin comigratory protein [Haemophilus influenzae]	64	41	162
603	4	759	956	gil 755823	NADH dehydrogenase F [Streptogyna americana]	64	35	198
653	2	940	746	gil 1213234	dicarboxylic amino acids Dip5p permease [Saccharomyces cerevisiae]	64	41	195
660	3	3801	2257	spi P46133 YDAH	HYPOTHETICAL PROTEIN IN OGT 5'REGION (FRAGMENT).	64	39	1545
695	1	11	502	gil 1001383	hypothetical protein [Synechocystis sp.]	64	41	492
702	1	3	752	gil 142865	DNA primase [Bacillus subtilis]	64	46	750
826	1	1	339	gil 971336	arginyl tRNA synthetase [Bacillus subtilis]	64	50	339
838	1	1831	917	gil 1354775	pfos/R [Treponema pallidum]	64	41	915
864	3	675	944	gil 39833	cyclomaltodextrin glucanotransferase [Bacillus stearothermophilus] i 39835	64	47	270
887	1	3	677	gil 153002	enterotoxin type E precursor [Staphylococcus aureus] pir A28179 A28179	64	46	675
					enterotoxin E precursor - Staphylococcus aureus sp P12993 ETXE_STAAU			
					ENTEROTOXIN TYPE E PRECURSOR (SEE).			
928	2	1172	963	gil 311976	fibrinogen-binding protein [Staphylococcus aureus] pir S34270 S34270	64	41	210
					fibrinogen-binding protein - Staphylococcus ureus			
1049	2	800	606	gil 1049115	Rap60 [Bacillus subtilis]	64	42	195
1067	2	999	748	gil 1151072	hHda precursor [Haemophilus ducreyi]	64	50	252

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1120	1	50	202	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	64	30	153
1125	1	751	377	gi 581648	lepB gene product [Staphylococcus epidermidis]	64	44	375
1688	1	402	214	pir A01365 TVMS	transforming protein K-ras - mouse	64	47	189
2472	1	2	358	gi 487282	Na+ -ATPase subunit J [Enterococcus hirae]	64	36	357
2989	1	520	356	gi 304134	larcC [Bacillus stearothermophilus]	64	50	165
3013	1	630	352	gi 551699	cytochrome oxidase subunit I [Bacillus firmus]	64	51	279
3034	1	546	274	gi 1204349	hypothetical protein (GB:D90212_3) [Haemophilus influenzae]	64	50	273
3197	1	613	308	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	64	46	306
3303	1	90	362	gi 1107839	alginate lyase [Pseudomonas aeruginosa]	64	43	273
3852	2	82	288	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	64	42	207
3868	1	1	312	gi 149435	putative [Lactococcus lactis]	64	48	312
3918	1	660	331	gi 5532	acetyl-CoA acyltransferase [Yarrowia lipolytica]	64	46	330
4000	1	112	378	gi 984688	unknown [Saccharomyces cerevisiae]	64	44	267
4009	1	81	368	gi 39372	grsB gene product [Bacillus brevis]	64	41	288
4166	1	2	349	gi 149435	putative [Lactococcus lactis]	64	46	348
4366	1	2	307	gi 216267	ORF2 [Bacillus megaterium]	64	44	306
4457	1	2	400	gi 1197667	vitellogenin [Anolis pulchellus]	64	43	399
11	3	1539	2438	gi 438228	ORF C [Staphylococcus aureus]	63	32	900
24	7	5611	5423	gi 1369943	a1 gene product [Bacteriophage B1]	63	34	189
29	1	1	390	gi 467441	expressed at the end of exponential growth under conditions in which he enzymes of the TCA cycle are repressed [Bacillus subtilis] gi 467441	63	43	390
					expressed at the end of exponential growth under onditions in which the enzymes of the TCA cycle are repressed Bacil			
31	6	6329	5712	gi 496943	ORF [Saccharomyces cerevisiae]	63	47	618
44	23	14669	15019	pir A04446 QDEC	hypothetical protein F-92 - Escherichia coli	63	36	351
48	6	4403	6250	gi 43498	pyruvate synthase [Halobacterium halobium]	63	42	1848
50	5	3869	4738	gi 413967	ipa-43d gene product [Bacillus subtilis]	63	43	870
53	6	6764	5742	gi 474176	regulator protein [Staphylococcus xylosus]	63	49	1023

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
56	114	115880	117607	gi 467409	DNA polymerase III subunit [Bacillus subtilis]	63	44	1728
57	111	7945	7376	gi 537036	ORF_o158 [Escherichia coli]	63	39	570
62	3	2479	2114	gi 642656	unknown [Rhizobium meliloti]	63	41	366
70	8	6562	7353	gi 1399821	PhoC [Rhizobium meliloti]	63	46	792
75	2	223	927	gi 149376	HisG [Lactococcus lactis]	63	45	705
78	5	4912	4403	gi 413950	ipa-26d gene product [Bacillus subtilis]	63	42	510
91	5	9076	7220	gi 466997	methH; B2126_C1_157 [Mycobacterium leprae]	63	41	1857
91	8	10566	9448	gi 1204344	cystathionine gamma-synthase [Haemophilus influenzae]	63	45	1119
120	1	21	1508	gi 882657	sulfite reductase (NADPH) flavoprotein beta subunit [Escherichia coli]	63	46	1488
120	4	2722	4125	gi 665994	hypothetical protein [Bacillus subtilis]	63	34	1404
127	7	6064	7566	gi 40162	murE gene product [Bacillus subtilis]	63	44	1503
149	6	2321	2106	gi 148503	dnak [Erysipelothrix rhusiopathiae]	63	40	216
149	26	10445	10170	gi 4870	ORF 2, has similarity to DNA polymerase [Saccharomyces kluyveri]	63	42	276
				gi 515961 S15961	hypothetical protein 2 - yeast [Saccharomyces yveri]			
					plasmid pSKL			
164	2	507	1298	gi 145476	CDP-diglyceride synthetase [Escherichia coli]	63	44	792
166	6	9909	8164	gi 151932	fructose enzyme II [Rhodobacter capsulatus]	63	41	1746
169	4	1704	1886	gi 152886	elongation factor Ts (tsf) [Spiroplasma citri]	63	48	183
188	5	3145	2951	gi 1334547	GIY COI i14 grp IB protein [Podospora anserina]	63	42	195
195	13	11767	12804	gi 606100	ORF_o335 [Escherichia coli]	63	40	1038
201	2	607	2283	gi 433534	arginyl-tRNA synthetase [Corynebacterium glutamicum] pir A49936 A49936	63	46	1677
					arginine--tRNA ligase (EC 6.1.1.19) - orynebacterium glutamicum			
206	14	15893	16489	gi 580828	N-acetyl-glutamate-gamma-semialdehyde dehydrogenase [Bacillus ubtilis]	63	49	597
220	5	7769	5766	gi 216334	seca protein [Bacillus subtilis]	63	42	2004
221	1	74	907	gi 677945	AppA [Bacillus subtilis]	63	42	834
227	3	944	1708	gi 1510558	icobyric acid synthase [Methanococcus jannaschii]	63	46	765
261	2	804	1070	gi 486511	ORF YKR054c [Saccharomyces cerevisiae]	63	45	267
269	2	3606	1960	gi 148221	DNA-dependent ATPase, DNA helicase [Escherichia coli] pir JS0137 BVECRQ	63	42	1647
					recQ protein - Escherichia coli			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
278	8	7417	6176	gi 699273	cystathionine gamma-synthase [Mycobacterium leprae] sp P46807 METB MYCLE	63	41	1242
					CYSTATHIONINE GAMMA-SYNTHASE (EC 4.2.99.9) O-SUCCINYLMOSERINE (THIOL)-LYASE).			
287	2	738	1733	gi 405133	putative [Bacillus subtilis]	63	38	996
295	1	2	748	gi 1239983	hypothetical protein [Bacillus subtilis]	63	41	747
328	3	2148	3134	gi 45302	carrier protein (AA 1 - 437) [Pseudomonas aeruginosa] ir S11497 S11497	63	36	987
					branched-chain amino acid transport protein braB - eudomonas aeruginosa			
362	2	1626	1216	sp P35136 SERA	D-3-PHOSPHOGLYCERATE DEHYDROGENASE (EC 1.1.1.95) (PGDH).	63	38	411
404	1	326	1051	gi 1303816	YgeZ [Bacillus subtilis]	63	35	726
405	3	2101	1715	gi 1303914	YqhY [Bacillus subtilis]	63	42	387
406	1	451	227	gi 142152	sulfate permease (gtg start codon) [Synechococcus PCC6301]	63	43	225
					pir A30301 GRYCS7 sulfate transport protein - Synechococcus sp. PCC 7942)			
415	2	1048	2718	gi 1205402	transport ATP-binding protein [Haemophilus influenzae]	63	41	1671
426	4	3575	2679	gi 393268	29-kilodalton protein [Streptococcus pneumoniae] sp P42362 P29K_STRPN 29 KD	63	39	897
					MEMBRANE PROTEIN IN PSAA 5'REGION ORF1).			
505	3	1347	2195	gi 1418999	orf4 [Lactobacillus sake]	63	40	849
507	1	2	574	gi 546917	comK [Bacillus subtilis, E26, Peptide, 192 aa]	63	35	573
562	2	146	1084	gi 43985	nifs-like gene [Lactobacillus delbrueckii]	63	45	939
675	1	427	215	gi 1510994	serine aminotransferase [Methanococcus jannaschii]	63	29	213
686	1	3	230	gi 517356	nitrate reductase (NADH) [Lotus japonicus]	63	52	228
701	1	3	392	gi 881940	NorQ protein [Paracoccus denitrificans]	63	41	390
720	1	2	400	gi 47168	open reading frame [Streptomyces lividans]	63	35	399
779	1	571	287	gi 1261932	unknown [Mycobacterium tuberculosis]	63	41	285
907	1	22	321	gi 149445	ORF1 [Lactococcus lactis]	63	27	300
972	1	794	399	gi 1511235	M. jannaschii predicted coding region MJ1232 [Methanococcus jannaschii]	63	27	396
1085	1	1154	618	gi 1204277	hypothetical protein (GB:U00019.14) [Haemophilus influenzae]	63	38	537
1094	1	3	542	gi 790943	urea amidolyase [Bacillus subtilis]	63	39	540
1108	1	3	482	pir S49892 S498	regulation protein - Bacillus subtilis	63	44	480
1113	1	1231	617	gi 493017	endocarditis specific antigen [Enterococcus faecalis]	63	45	615

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1300	1	3	695	sp P33940 YOHJ	HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION.	63	46	693
1325	1	1	204	gi 928989	p100 protein [Borrelia burgdorferi]	63	30	204
1814	1	3	245	gi 1303914	YqhY [Bacillus subtilis]	63	34	243
2021	1	498	250	pir C33496 C334	hisc homolog - Bacillus subtilis	63	46	249
2325	1	2	193	gi 436132	product is similar to TnpA of transposon Tn554 from Staphylococcus aureus [Clostridium butyricum]	63	40	192
2335	1	1	195	gi 1184298	flagellar MS-ring protein [Borrelia burgdorferi]	63	47	195
2406	1	451	227	gi 1041785	rhoptyr protein [Plasmodium yoelii]	63	33	225
2961	2	136	360	gi 312443	carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus aldolyticus]	63	52	225
2965	1	1	402	gi 1407784	orf-1; novel antigen [Staphylococcus aureus]	63	50	402
2987	1	583	293	gi 1224069	amidase [Moraxella catarrhalis]	63	35	291
2994	1	266	135	gi 836646	phosphoribosylformimino-praic ketoisomerase [Rhodobacter phaeroides]	63	51	132
3043	1	440	252	gi 1480237	phenylacetaldehyde dehydrogenase [Escherichia coli]	63	40	189
3078	1	609	400	gi 1487982	intrinsic membrane protein [Mycoplasma hominis]	63	36	210
3139	1	2	217	gi 439126	glutamate synthase (NADPH) [Azospirillum brasilense] pir A49916 A49916	63	47	216
3625	1	793	398	gi 623073	ORF360; putative [Bacteriophage LL-H]	63	48	396
3658	1	1	399	gi 1303697	YrxA [Bacillus subtilis]	63	37	399
3659	1	3	395	gi 1256135	YbbF [Bacillus subtilis]	63	48	393
3783	1	720	361	gi 1256902	Pyruvate decarboxylase isozyme 2 (Swiss Prot. accession number P16467) [Saccharomyces cerevisiae]	63	34	360
3900	1	338	171	sp P10537 AMYB	BETA-AMYLASE (EC 3.2.1.2) (1,4-ALPHA-D-GLUCAN MALTOHYDROLASE).	63	54	168
4309	1	3	176	pir A37967 A379	neural cell adhesion molecule Ng-CAM precursor - chicken	63	57	174
4367	1	1	195	gi 1321932	Per6p gene product [Pichia pastoris]	63	30	195
4432	1	1	312	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756	63	51	312
4468	1	6	308	gi 296464	ATPase [Lactococcus lactis]	63	36	303
33	3	1411	2400	gi 153675	tagatose 6-P kinase [Streptococcus mutans]	62	44	990
36	9	5985	6218	gi 1490521	hMSH3 [Homo sapiens]	62	51	234

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
37	1	2	721	gi 1107531	ceuE gene product [Campylobacter coli]	62	33	720
38	15	10912	11589	gi 1222058	H. influenzae predicted coding region HIN1279 [Haemophilus influenzae]	62	38	678
38	25	119526	20329	gi 695280	ORF2 [Alcaligenes eutrophus]	62	41	804
57	2	2523	1780	gi 471234	orf1 [Haemophilus influenzae]	62	55	744
57	9	6646	6350	gi 508174	EIIB domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	62	35	297
58	1	2	559	gi 755152	highly hydrophobic integral membrane protein [Bacillus subtilis] sp P42953 TAGG_BACSU TEICHOIC ACID TRANSLOCATION PERMEASE PROTEIN AGG.	62	34	558
67	10	8250	9014	gi 470683	Shows similarity with ATP-binding proteins from other ABC-transport perons, Swiss Prot Accession Numbers P24137, P08007, P04285, P24136 Escherichia coli	62	34	765
69	8	8315	7494	gi 46816	actVA 4 gene product [Streptomyces coelicolor]	62	44	822
80	3	1793	1320	gi 39993	UDP-N-acetylmuramoylalanine--D-glutamate ligase [Bacillus subtilis]	62	43	474
87	7	7034	9205	gi 217191	5'-nucleotidase precursor [Vibrio parahaemolyticus]	62	48	2172
100	3	4051	3089	gi 1511047	phosphoglycerate dehydrogenase [Methanococcus jannaschii]	62	42	963
102	1	2	520	gi 153655	mismatch repair protein [Streptococcus pneumoniae] pir C28667 C28667 DNA mismatch repair protein hexA - Streptococcus pneumoniae	62	34	519
112	2	466	1068	gi 153741	ATP-binding protein [Streptococcus mutans]	62	37	603
114	7	6855	7562	gi 1204866	L-fucose operon activator [Haemophilus influenzae]	62	38	708
116	4	6823	5633	gi 677947	AppC [Bacillus subtilis]	62	37	1191
124	8	6855	6004	gi 853777	product similar to E.coli PRFA2 protein [Bacillus subtilis] pir S55438 S55438 ywke protein - Bacillus subtilis sp P45873 HEMK_BACSU POSSIBLE PROTOPORPHYRINOGEN OXIDASE (EC 3.3.-).	62	44	852
148	1	24	554	gi 467456	unknown [Bacillus subtilis]	62	50	531
149	20	7591	6725	gi 1205807	replicative DNA helicase [Haemophilus influenzae]	62	41	867
163	3	1503	1153	gi 40067	X gene product [Bacillus sphaericus]	62	42	351
164	15	14673	15632	gi 42219	P35 gene product (AA 1 - 314) [Escherichia coli]	62	38	960
165	2	1166	1447	gi 403936	phenylalanyl-tRNA synthetase alpha subunit (Gly294 variant) unidentified cloning vector	62	38	282
166	2	2084	5089	gi 308861	GTG start codon [Lactococcus lactis]	62	44	3006
171	1	1225	614	gi 1046053	hypothetical protein (SP:P32049) [Mycoplasma genitalium]	62	41	612

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
183	1	2521	1310	gi 143045	hemY [Bacillus subtilis]	62	45	1212
200	1	3	956	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	62	32	954
237	2	935	1966	gi 41695	hisc protein [Escherichia coli]	62	44	1032
261	3	4008	2605	gi 143121	ORF A; putative [Bacillus firmus]	62	42	1404
299	8	4477	4719	gi 467441	expressed at the end of exponential growth under conditions in which he enzymes of the TCA cycle are repressed [Bacillus subtilis] gi 467441	62	47	243
					expressed at the end of exponential growth under ondtions in which the enzymes of the TCA cycle are repressed Bacil			
304	6	5018	3819	gi 153015	FemA protein [Staphylococcus aureus]	62	43	1200
324	1	2	262	gi 142717	cytochrome aa3 controlling protein [Bacillus subtilis] pir A33960 A33960	62	30	261
					cta protein - Bacillus subtilis sp P12946 CTAA_BACSU CYTOCHROME AA3			
					CONTROLLING PROTEIN.			
325	2	269	1207	gi 581088	methionyl-tRNA formyltransferase [Escherichia coli]	62	39	939
332	6	4894	4631	gi 1499960	uridine 5'-monophosphate synthase [Methanococcus jannaschii]	62	36	264
355	1	2	370	gi 145925	fecB [Escherichia coli]	62	32	369
365	8	6628	6804	gi 413943	ipa-19d gene product [Bacillus subtilis]	62	54	177
369	2	2744	1626	pir A43577 A435	regulatory protein pfor - Clostridium perfringens	62	42	1119
370	1	34	264	gi 40665	beta-glucosidase [Clostridium thermoceillum]	62	37	231
415	3	2709	3176	gi 1205401	transport ATP-binding protein [Haemophilus influenzae]	62	35	468
429	1	1578	790	gi 1046024	[Na+ ATPase subunit J [Mycoplasma genitalium]	62	40	789
444	2	704	1369	gi 581510	nodulation gene; integral membrane protein; homology to Rhizobium eguminosarum nodI [Rhizobium loti]	62	37	666
477	2	751	1869	pir A48440 A484	ring-infected erythrocyte surface antigen 2, RESA-2 - Plasmodium falciparum	62	44	1119
485	1	241	1707	gi 17934	betaine aldehyd dehydrogenase [Beta vulgaris]	62	43	1467
487	3	1141	1311	gi 149445	ORF1 [Lactococcus lactis]	62	31	171
494	2	1134	1313	gi 166835	ribulose bisphosphate carboxylase/oxygenase activase [Arabidopsis haliana]	62	37	180
518	1	193	882	gi 153491	O-methyltransferase [Streptomyces glaucescens]	62	39	690
534	2	369	2522	gi 1480429	putative transcriptional regulator [Bacillus steaothermophilus]	62	35	2154
551	6	4371	4820	gi 511113	feric uptake regulation protein [Campylobacter jejuni]	62	37	450
574	1	1	570	gi 153000	enterotoxin B [Staphylococcus aureus]	62	43	570

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
590	2	344	1171	gi 40367	ORFC [Clostridium acetobutylicum]	62	37	828
655	1	396	830	gi 147195	phnB protein [Escherichia coli]	62	44	435
656	1	2	478	gi 1205451	cell division inhibitor [Haemophilus influenzae]	62	36	477
676	1	692	348	gi 1511613	[methyl coenzyme M reductase system, component A2 [Methanococcus jannaschii]]	62	36	345
687	1	493	248	gi 49272	Asparaginase [Bacillus licheniformis]	62	48	246
700	2	267	944	gi 1205822	hypothetical protein (GB:X75627.4) [Haemophilus influenzae]	62	40	678
840	2	1715	1041	gi 1045865	[M. genitalium predicted coding region MG181 [Mycoplasma genitalium]]	62	36	675
864	4	898	1491	gi 1144332	deoxyuridine nucleotidohydrolase [Homo sapiens]	62	38	594
916	1	35	400	gi 413931	ipa-7d gene product [Bacillus subtilis]	62	45	366
1071	1	1	771	gi 1510649	aspartokinase I [Methanococcus jannaschii]	62	40	771
1084	1	19	609	gi 688011	AgX-1 antigen [human, infertile patient, testis, Peptide, 505 aa]	62	39	591
1103	1	3	203	gi 581261	[ORF homologous to E.coli metB [Herpetosiphon aurantiacus] pir S14030 S14030] Hypothetical protein - Herpetosiphon aurantiacus fragment)	62	51	201
1217	1	463	233	gi 460025	ORF2, putative [Streptococcus pneumoniae]	62	41	231
1533	1	644	414	gi 413968	ipa-44d gene product [Bacillus subtilis]	62	48	231
1537	1	3	257	gi 1510641	[alanyl-tRNA synthetase [Methanococcus jannaschii]]	62	29	255
2287	1	3	161	gi 485956	mrpC gene product [Proteus mirabilis]	62	45	159
2386	1	3	245	gi 285708	[nontoxic component [Clostridium botulinum]]	62	31	243
2484	1	331	167	gi 142092	[DNA-repair protein (recA) [Anabaena variabilis]]	62	35	165
2490	1	798	400	gi 581648	epiB gene product [Staphylococcus epidermidis]	62	42	399
3016	1	596	300	gi 710022	uroporphyrinogen III [Bacillus subtilis]	62	51	297
3116	1	1	213	gi 466883	[nifs; B1496_C2_193 [Mycobacterium leprae]]	62	44	213
3297	1	823	413	gi 475715	[acetyl coenzyme A acetyltransferase (thiolase) [Clostridium acetobutylicum]]	62	42	411
3609	1	31	276	gi 1408501	[homologous to N-acyl-L-amino acid amidohydrolase of Bacillus stearothermophilus [Bacillus subtilis]]	62	48	246
3665	2	584	402	gi 151259	[HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.]	62	40	183
3733	1	3	374	gi 1353197	[thioredoxin reductase [Eubacterium acidaminophilum]]	62	42	372

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3898	1	1	237	gi 153675	tagatose 6-P kinase [Streptococcus mutans]	62	45	237
4027	1	283	143	gi 330705	homologue to gene 30 (aa 1-59); putative [Bovine herpesvirus 4]	62	43	141
4109	1	727	365	gi 41748	hsdM protein (AA 1-520) [Escherichia coli]	62	45	363
4303	1	1	303	gi 1303813	YqgW [Bacillus subtilis]	62	43	303
4380	1	530	267	gi 1235684	mevalonate pyrophosphate decarboxylase [Saccharomyces cerevisiae]	62	55	264
4494	1	2	256	gi 510692	enterotoxin H [Staphylococcus aureus]	62	34	255
4598	1	411	223	gi 763513	ORF4; putative [Streptomyces violaceoruber]	62	45	189
4624	1	1	222	gi 41748	hsdM protein (AA 1-520) [Escherichia coli]	62	45	222
5	5	4288	3932	gi 928831	ORF95; putative [Lactococcus lactis phage BK5-T]	61	36	357
11	1	320	162	pir C33356 C333	prothymosin alpha homolog (clone 32) - human (fragment)	61	33	159
16	11	10991	11938	gi 1205391	hypothetical protein (SP:P33995) [Haemophilus influenzae]	61	44	948
32	1	283	801	gi 1066504	exo-beta 1,3 glucanase [Cochliobolus carbonum]	61	50	519
38	3	616	1107	gi 1510864	glutamine transport ATP-binding protein Q [Methanococcus jannaschii]	61	41	492
45	4	3082	4038	gi 1109686	ProX [Bacillus subtilis]	61	45	957
48	8	7118	7504	gi 498839	ORF2 [Clostridium perfringens]	61	33	387
51	9	4605	5570	gi 388269	traC [Plasmid pAD1]	61	42	966
60	6	1689	2243	gi 1205893	hypothetical protein (GB:U00011.3) [Haemophilus influenzae]	61	32	555
62	9	5559	5122	gi 854656	Na/H antiporter system ORF2 [Bacillus alcalophilus]	61	38	438
67	5	4330	5646	gi 466612	nika [Escherichia coli]	61	36	1317
74	2	2400	1504	gi 1204846	carbamate kinase [Haemophilus influenzae]	61	40	897
85	1	2198	1101	gi 1498756	amidophosphoribosyltransferase PurF [Rhizobium etli]	61	41	1098
86	4	1995	1582	gi 1499931	M. jannaschii predicted coding region MJ1083 [Methanococcus jannaschii]	61	44	414
97	1	74	649	gi 1518679	orf [Bacillus subtilis]	61	44	576
99	2	2454	1990	gi 413958	ipa-34d gene product [Bacillus subtilis]	61	18	465
124	7	6223	5123	gi 556881	Similar to Saccharomyces cerevisiae SUA5 protein [Bacillus subtilis] pir S49358 S49358 ipc-29d protein - Bacillus subtilis sp P39153 YWLC_BACSU HYPOTHETICAL 37.0 KD PROTEIN IN SPOIIR-GLYC NTERGENIC REGION.	61	46	1101
125	4	1668	2531	gi 1491643	ORF4 gene product [Chloroflexus aurantiacus]	61	43	864

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
132	1	1250	627	pir PQ0259 PQ02	hypothetical protein I - Enterococcus faecalis plasmid pAM-beta-1 (fragment)	61	43	624
149	9	3617	3075	gi 1144332	deoxyuridine nucleotidohydrolase [Homo sapiens]	61	40	543
149	22	8690	7869	gi 160047	p101/acidic basic repeat antigen [Plasmodium falciparum] pir A29232 A29232 101K malaria antigen precursor - Plasmodium alciparum (strain Camp)	61	35	822
168	3	1915	2361	gi 1499694	HIT protein, member of the HIT-family [Methanococcus jannaschii]	61	41	447
171	9	9675	7948	gi 467446	similar to SpoVB [Bacillus subtilis]	61	38	1728
174	3	1042	2340	gi 216374	glutaryl 7-ACA acylase precursor [Bacillus laterosporus]	61	49	1299
190	4	5034	4111	gi 409286	bmrU [Bacillus subtilis]	61	37	924
216	1	2	190	gi 415861	eukaryotic initiation factor 2 beta (eIF-2 beta) [Oryctolagus unicusulus]	61	29	189
227	7	4161	5048	gi 216341	ORF for methionine amino peptidase [Bacillus subtilis]	61	41	888
238	4	1959	3047	gi 809543	CbrC protein [Erwinia chrysanthemi]	61	38	1089
247	1	2	694	gi 537231	ORF_f579 [Escherichia coli]	61	38	693
247	2	678	1034	gi 142226	chvD protein [Agrobacterium tumefaciens]	61	40	357
257	2	3523	2627	gi 699379	glvr-1 protein [Mycobacterium leprae]	61	40	897
268	2	3419	3051	gi 40364	ORFAl [Clostridium acetobutylicum]	61	41	369
275	4	4621	4827	gi 1204848	hypothetical protein (GP:M87049.57) [Haemophilus influenzae]	61	36	207
277	1	1	1845	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	45	1845
278	9	8003	7032	gi 467462	cysteine synthetase A [Bacillus subtilis]	61	43	972
278	10	9878	8535	gi 1205919	Na+ and Cl- dependent gamma-aminobutyric acid transporter [Haemophilus influenzae]	61	38	1344
283	1	1	366	gi 755607	polyA polymerase [Bacillus subtilis]	61	36	366
288	2	1918	1496	gi 388108	cell wall enzyme [Enterococcus faecalis]	61	43	423
291	1	86	334	gi 454265	FBP3 [Petunia hybrida]	61	38	249
318	1	1104	694	gi 290531	similar to beta-glucoside transport protein [Escherichia coli] sp P31451 PTIB ECOLI PTS SYSTEM, ARBUTIN-LIKE IIB COMPONENT PHOSPHOTRANSFERASE ENZYME II, B COMPONENT (EC 2.7.1.69).	61	47	411
330	2	1912	1190	gi 1001805	hypothetical protein [Synechocystis sp.]	61	41	723

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
385	2	1513	1025	gi1533098	DnaD protein [Bacillus subtilis]	61	42	489
426	1	794	399	gi11303853	YqgF [Bacillus subtilis]	61	44	396
438	3	810	1421	gi1293660	AbsA2 [Streptomyces coelicolor]	61	36	612
454	1	1580	792	gi1733522	phosphatidylinositol-4,5-diphosphate 3-kinase [Dictyostelium discoideum]	61	30	789
464	2	784	560	gi11123120	C53B7.5 gene product [Caenorhabditis elegans]	61	38	225
470	8	6077	7357	gi1623073	ORF360; putative [Bacteriophage LL-H]	61	47	1281
509	1	554	279	gi1467484	unknown [Bacillus subtilis]	61	45	276
555	3	1916	1296	gi1141800	anthranilate synthase glutamine amidotransferase [Acinetobacter alcoaceticus]	61	42	621
569	1	1711	857	gi1467090	B2235_C2_195 [Mycobacterium leprae]	61	47	855
585	2	961	803	spIP36686 SURE_	SURVIVAL PROTEIN SURE HOMOLOG (FRAGMENT) .	61	33	159
592	3	1694	1422	gi1221602	immunity repressor protein [Haemophilus influenzae]	61	32	273
603	1	43	357	gi1507738	Hmp [Vibrio parahaemolyticus]	61	33	315
669	1	2467	1235	gi11146243	22.4% identity with Escherichia coli DNA-damage inducible protein ...; putative [Bacillus subtilis]	61	37	1233
675	3	805	1101	gi1403373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis]	61	36	297
				pir S37251 S37251	glycerophosphoryl diester phosphodiesterase - acillus subtilis			
703	1	1656	829	gi1537181	ORF_f470 [Escherichia coli]	61	32	828
728	1	1628	816	gi1806281	DNA polymerase I [Bacillus stearothermophilus]	61	39	813
821	1	61	318	gi1709992	hypothetical protein [Bacillus subtilis]	61	38	258
856	2	2313	1567	gi1609310	portal protein gp3 [Bacteriophage HK97]	61	40	747
923	1	1081	542	gi1143213	putative [Bacillus subtilis]	61	38	540
1124	1	59	370	gi11107541	C33D9.8 [Caenorhabditis elegans]	61	26	312
1492	1	548	276	gi1406397	unknown [Mycoplasma genitalium]	61	32	273
1602	1	46	318	gi1733522	phosphatidylinositol-4,5-diphosphate 3-kinase [Dictyostelium discoideum]	61	34	273
2500	1	577	290	gi11045964	hypothetical protein (GB:U14003_297) [Mycoplasma genitalium]	61	31	288
2968	1	2	808	gi1397526	clumping factor [Staphylococcus aureus]	61	55	807
3076	1	3	248	gi1149373	ORF_1 [Lactococcus lactis]	61	41	246

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3609	2	207	401	gi 1408501	homologous to N-acyl-L-amino acid amidohydrolase of Bacillus stearothermophilus [Bacillus subtilis]	61	39	195
3662	1	1477	740	gi 1303813	Yqew [Bacillus subtilis]	61	42	738
3672	1	2	442	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	441
3724	1	2	220	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	61	41	219
3728	1	3	398	gi 677943	AppD [Bacillus subtilis]	61	46	396
3884	1	3	401	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	47	399
3971	1	3	383	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	45	381
4038	1	661	359	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	61	24	303
4041	1	546	274	gi 413953	ipa-29d gene product [Bacillus subtilis]	61	48	273
4047	1	1	402	gi 528991	unknown [Bacillus subtilis]	61	42	402
4102	1	1	345	gi 976025	HrsA [Escherichia coli]	61	46	345
4155	1	1	336	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	336
4268	1	463	233	gi 450688	hsdM gene of EcoprrI gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	61	38	231
4374	1	542	273	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	270
4389	1	2	172	gi 147516	ribokinase [Escherichia coli]	61	35	171
4621	1	2	268	gi 784897	beta-N-acetylhexosaminidase [Streptococcus pneumoniae] pir A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	47	267
4663	1	27	227	gi 976025	HrsA [Escherichia coli]	61	50	201
4	6	6663	5536	gi 1408501	homologous to N-acyl-L-amino acid amidohydrolase of Bacillus stearothermophilus [Bacillus subtilis]	60	43	1128

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
11	6	3426	3725	gi 410748	ring-infested erythrocyte surface antigen [Plasmodium falciparum]	60	24	300
					pir A25526 A25526 ring-infected erythrocyte surface antigen precursor - Plasmodium falciparum (strain FC27/Papua New Guinea) sp P13830 RESA_PLAFF			
					RING-INFECTED ERYTHROCYTE SURFACE ANTIGEN RE			
11	14	11035	10313	gi 1217651	carbonyl reductase (NADPH) [Rattus norvegicus]	60	28	723
16	12	11917	12930	gi 1001453	hypothetical protein [Synechocystis sp.]	60	37	1014
33	1	26	469	gi 388109	regulatory protein [Enterococcus faecalis]	60	41	444
37	13	110814	9834	gi 1336656	Orf1 [Bacillus subtilis]	60	40	981
39	4	4364	4522	gi 4872	ORF 4 [Saccharomyces kluyveri]	60	47	159
41	1	2047	1025	gi 142822	D-alanine racemase cds [Bacillus subtilis]	60	39	1023
43	4	2474	3607	gi 468046	para-nitrobenzyl esterase [Bacillus subtilis]	60	40	1134
44	10	6756	7769	gi 414234	thiF [Escherichia coli]	60	52	1014
45	10	8874	9074	gi 343949	var1(40.0) [Saccharomyces cerevisiae]	60	44	201
56	18	127842	126430	gi 468764	mocR gene product [Rhizobium meliloti]	60	35	1413
60	2	173	388	gi 1303864	YqgQ [Bacillus subtilis]	60	33	216
63	2	357	1619	gi 467124	ureD; B229_C3_234 [Mycobacterium leprae]	60	43	1263
69	1	787	395	gi 1518853	OafA [Salmonella typhimurium]	60	36	393
88	1	1	1188	gi 1480429	putative transcriptional regulator [Bacillus stearothermophilus]	60	30	1188
92	6	4735	3881	gi 349227	transmembrane protein [Escherichia coli]	60	37	855
92	7	5996	4923	gi 466613	nikB [Escherichia coli]	60	38	1074
93	1	949	476	gi 1510925	coenzyme F420-reducing hydrogenase, beta subunit [Methanococcus jannaschii]	60	27	474
96	6	7366	7578	gi 972715	accessory protein [Carnobacterium piscicola]	60	30	213
98	6	3212	4069	gi 467425	unknown [Bacillus subtilis]	60	42	858
102	10	7158	7430	gi 143092	acetolactate synthase small subunit [Bacillus subtilis]	60	37	273
					sp P37252 ILVN_BAGSU ACETOLACTATE SYNTHASE SMALL SUBUNIT (EC .1.3.18) (AHAS) (ACETOHYDROXY-ACID SYNTHASE SMALL SUBUNIT) (ALS).			
109	11	9127	10515	gi 1255259	o-succinylbenzoic acid (OSB) CoA ligase [Staphylococcus aureus]	60	28	1389
109	12	110499	11656	gi 141954	beta-ketothiolase [Alcaligenes eutrophus]	60	41	1158
119	2	4630	3134	gi 1524280	unknown [Mycobacterium tuberculosis]	60	45	1497

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
121	9	6957	7646	gi 1107529	ceuC gene product [Campylobacter coli]	60	35	690
140	7	7704	6013	gi 146547	kbpA [Escherichia coli]	60	45	1692
145	1	2	703	gi 1460077	unknown [Mycobacterium tuberculosis]	60	23	702
150	3	2809	2216	gi 1146230	putative [Bacillus subtilis]	60	40	594
157	2	1389	961	gi 1303975	YqjX [Bacillus subtilis]	60	30	429
158	5	5125	4769	gi 1449288	unknown [Mycobacterium tuberculosis]	60	36	357
159	1	511	257	gi 580932	murD gene product [Bacillus subtilis]	60	43	255
160	1	159	1187	gi 1204532	hypothetical protein (GB:U19201_29) [Haemophilus influenzae]	60	34	1029
161	14	8249	7866	gi 1496003	ORF3; PepY; putative oligoendopeptidase based on homology with Lactococcus lactis PepF (GenBank Accession Number Z32522) [Caldicellulosiruptor saccharolyticus]	60	34	384
172	3	1331	2110	gi 485280	28.2 kDa protein [Streptococcus pneumoniae]	60	33	780
173	2	4082	2460	gi 1524397	glycine betaine transporter OpuD [Bacillus subtilis]	60	41	1623
173	4	5963	4953	gi 1100737	NADP dependent leukotriene b4 12-hydroxydehydrogenase [Sus scrofa]	60	44	1011
198	1	3	995	gi 413943	ipa-19d gene product [Bacillus subtilis]	60	42	993
201	4	3641	4573	sp P37028 YADF_HYPOTHETICAL 29.4 KD PROTEIN IN HEML-PFS INTERGENIC REGION PRECURSOR.		60	37	933
203	3	3269	2415	gi 927798	D9719.34p; CAI: 0.14 [Saccharomyces cerevisiae]	60	43	855
206	9	12234	12515	sp P37347 YECD_HYPOTHETICAL 21.8 KD PROTEIN IN ASP5 5'REGION.		60	47	282
212	4	1213	1410	gi 332711	hemagglutinin-neuraminidase fusion protein [Human parainfluenza irus 3]	60	34	198
214	1	65	1153	gi 1204366	hypothetical protein (GB:U14003_130) [Haemophilus influenzae]	60	36	1089
237	1	2	937	gi 149377	HisD [Lactococcus lactis]	60	40	936
241	6	5696	4998	gi 1046160	hypothetical protein (GB:U00021_5) [Mycoplasma genitalium]	60	37	699
260	6	5919	6485	gi 431950	similar to a B.subtilis gene (GB: BACHEMEHY_5) [Clostridium asteurianum]	60	35	567
264	1	2432	1218	gi 397526	clumping factor [Staphylococcus aureus]	60	53	1215
267	1	3	1409	gi 148316	Nah-antiporter protein [Enterococcus hirae]	60	27	1407
275	3	3804	4595	pir F36889 F368	leuD 3'-region hypothetical protein - Lactococcus lactis subsp. lactis (strain IL1403)	60	35	792
291	3	860	1198	gi 1208889	coded for by C. elegans cDNA yk130e12.5; contains C2H2-type zinc fingers [Caenorhabditis elegans]	60	33	339

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
307	6	3421	3176	gi 1070014	protein-dependent [Bacillus subtilis]	60	36	246
316	8	4957	5823	gi 413952	ipa-28d gene product [Bacillus subtilis]	60	41	867
328	4	2996	3484	gi 1204484	membrane-associated component, branched amino acid transport system [Haemophilus influenzae]	60	39	489
332	5	4887	4363	gi 1205449	colicin V production protein (pur regulon) [Haemophilus influenzae]	60	37	525
357	1	1062	532	gi 887842	single-stranded DNA-specific exonuclease [Escherichia coli]	60	41	531
375	2	96	362	gi 4857	adenylyl cyclase gene product [Saccharomyces kluyveri] r JQ1145 OYBYK	60	47	267
397	1	66	416	gi 709999	Glucarate dehydratase [Bacillus subtilis]	60	37	351
409	1	2	163	gi 499700	glycogen phosphorylase [Saccharomyces cerevisiae]	60	35	162
453	4	914	1237	gi 1196899	unknown protein [Staphylococcus aureus]	60	36	324
453	7	3838	3620	sp P12222 YCF1_HYPOTHETICAL 226 KD PROTEIN (ORF 1901).		60	31	219
470	2	622	945	pir S30782 S307	integrin homolog - yeast (Saccharomyces cerevisiae)	60	31	324
500	1	118	606	gi 467407	unknown [Bacillus subtilis]	60	36	489
503	3	752	982	gi 167835	myosin heavy chain [Dictyostelium discoideum]	60	34	231
505	4	2238	3563	gi 1510732	NADH oxidase [Methanococcus jannaschii]	60	26	1326
523	1	3	1043	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis]	60	41	1041
				pir A27650 A27650	regulatory protein phoR - Bacillus subtilis			
				sp P23545 PHOR_BACSU	ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOR (EC 2.7.3.-)			
543	1	1	465	gi 1511103	cobalt transport ATP-binding protein O [Methanococcus jannaschii]	60	40	465
545	1	1	726	gi 1498192	putative [Pseudomonas aeruginosa]	60	40	726
556	1	2	1054	gi 1477402	tex gene product [Bordetella pertussis]	60	42	1053
578	1	974	489	gi 1205129	H. influenzae predicted coding region HI0882 [Haemophilus influenzae]	60	42	486
594	1	1	624	gi 1212755	adenylyl cyclase [Aeromonas hydrophila]	60	45	624
604	1	3	530	gi 145925	fecB [Escherichia coli]	60	42	528
620	1	926	465	gi 1205483	bicyclomycin resistance protein [Haemophilus influenzae]	60	33	462
630	2	871	1122	gi 1486242	unknown [Bacillus subtilis]	60	41	252
645	2	574	425	gi 1205136	serine hydroxymethyltransferase (serine methylase) [Haemophilus influenzae]	60	28	150

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
684	1	1082	843	gi 1205538	hypothetical protein (GB:U14003_302) [Haemophilus influenzae]	60	39	240
786	1	967	485	gi 1402944	orfM1 gene product [Bacillus subtilis]	60	46	483
844	1	588	346	gi 790943	urea amidolyase [Bacillus subtilis]	60	40	243
851	1	1	726	gi 159661	GMP reductase [Ascaris lumbricoides]	60	41	726
871	1	1746	874	gi 1001493	hypothetical protein [Synecocystis sp.]	60	39	873
896	1	1558	839	gi 604926	NADH dehydrogenase, subunit 5 [Schizopyllum commune] sp P50368 NU5M_SCHCO NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC .6.5.3).	60	39	720
908	2	448	753	gi 662880	novel hemolytic factor [Bacillus cereus]	60	31	306
979	1	2	595	gi 1429255	putative; orf1 [Bacillus subtilis]	60	30	594
1078	1	669	502	gi 581055	inner membrane copper tolerance protein [Escherichia coli] gi 871029 disulphide isomerase like protein [Escherichia coli] pir S47295 S47295 inner membrane copper tolerance protein - scherichia coli	60	40	168
1112	1	1150	620	gi 407885	ORF3 [Streptomyces griseus]	60	34	531
1135	1	484	275	gi 1171407	Vps8p [Saccharomyces cerevisiae]	60	36	210
1146	1	17	562	gi 1239981	hypothetical protein [Bacillus subtilis]	60	36	546
1291	1	716	360	pir S57530 S575	carboxyl esterase - Acinetobacter calcoaceticus	60	30	357
1332	1	336	169	gi 1222056	aminotransferase [Haemophilus influenzae]	60	44	168
1429	1	3	146	gi 1205619	ferritin like protein [Haemophilus influenzae]	60	39	144
1722	1	570	286	gi 240052	dihydroflavonol-4-reductase, DFR [Hordeum vulgare=barley, cv. Gula, eptide, 354 aa]	60	36	285
2350	1	385	200	gi 497626	ORF 1 [Plasmid pAQ1]	60	20	186
2936	1	519	310	gi 508981	prephenate dehydratase [Bacillus subtilis]	60	48	210
3027	1	568	302	gi 1146199	putative [Bacillus subtilis]	60	37	267
3084	1	20	208	gi 1407784	orf-1; novel antigen [Staphylococcus aureus]	60	51	189
3155	1	2	226	gi 1046097	cytadherence-accessory protein [Mycoplasma genitalium]	60	34	225
3603	1	368	186	gi 510108	mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA enhydrogenase alpha-subunit [Rattus norvegicus]	60	42	183
3665	1	486	244	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	60	42	243
3747	1	3	146	gi 474192	lucC gene product [Escherichia coli]	60	36	144

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3912	1	3	335	gi 1488695	[novel antigen; orf-2 [Staphylococcus aureus]	60	44	333
4072	1	3	272	gi 405879	[yeiH [Escherichia coli]	60	33	270
4134	1	510	352	gi 780656	[chemoreceptor protein [Rhizobium leguminosarum bv. viciae] gi 780656 [chemoreceptor protein [Rhizobium leguminosarum bv. iciae]	60	28	159
4207	2	677	402	gi 602031	[similar to trimethylamine DH [Mycoplasma capricolum] pir S49950 S49950 [probable trimethylamine dehydrogenase (EC .5.99.7) - Mycoplasma capricolum [SGC3] (fragment)]	60	41	276
4243	1	127	324	gi 899317	[peptide synthetase module [Microcystis aeruginosa] pir S49111 S49111 [probable amino acid activating domain - icrocystis aeruginosa (fragment) [SUB 144-528)]	60	42	198
4310	1	624	313	gi 508980	[pheB [Bacillus subtilis]	60	28	312
4345	1	343	173	gi 510108	[mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA ehydrogenase [alpha-subunit [Rattus norvegicus]	60	42	171
4382	1	498	280	gi 47382	[acyl-CoA-dehydrogenase [Streptomyces purpurascens]	60	48	219
4474	1	53	223	gi 510108	[mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA ehydrogenase [alpha-subunit [Rattus norvegicus]	60	42	171
23	4	4518	3523	gi 426446	[VipB protein [Salmonella typhi]	59	39	996
33	2	707	1483	pir S48604 S486	[hypothetical protein - Mycoplasma capricolum (SGC3) (fragment)]	59	33	777
33	5	4651	5853	gi 6721	[F59B2.3 [Caenorhabditis elegans]	59	33	1203
37	2	3228	2299	gi 142833	[ORF2 [Bacillus subtilis]	59	37	930
38	121	16784	16593	gi 912576	[BiP [Phaeodactylum tricornutum]	59	40	192
52	3	2648	2349	gi 536972	[ORF_o90a [Escherichia coli]	59	44	300
54	112	114181	13402	gi 483940	[transcription regulator [Bacillus subtilis]	59	37	780
57	3	4397	3339	gi 508176	[Gat-1-P-DH, NAD dependent [Escherichia coli]	59	40	1059
66	1	986	495	gi 1303901	[YqjT [Bacillus subtilis]	59	34	492
67	7	6552	7460	gi 912461	[nikC [Escherichia coli]	59	37	909
70	7	5383	6366	gi 1399822	[PhoD precursor [Rhizobium meliloti]	59	46	984
78	1	1	1449	gi 971345	[unknown, similar to E.coli cardiolipin synthase [Bacillus subtilis] [sp P45860 YWIE_BACSU HYPOTHETICAL 58.2 PROTEIN IN NARI-ACDA NTERGENIC [REGION.	59	39	1449
82	110	114329	15534	gi 490328	[LORF F [unidentified]	59	44	1206

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
89	2	1602	958	gi 642801	unknown [Saccharomyces cerevisiae]	59	32	645
96	4	4940	5473	gi 1333802	protein of unknown function [Rhodobacter capsulatus]	59	33	534
98	1	2	820	gi 467421	similar to B. subtilis DnaH [Bacillus subtilis]	59	34	819
119	1	166	1557	gi 143122	ORF B; putative [Bacillus firmus]	59	36	1392
120	10	6214	6756	gi 15354	ORF 55.9 [Bacteriophage T4]	59	39	543
120	16	12476	13510	gi 1086575	Beta [Rhizobium meliloti]	59	44	1035
123	1	386	195	gi 984737	catalase [Campylobacter jejuni]	59	38	192
130	1	370	645	gi 1256634	25.8% identity over 120 aa with the Synenococcus sp. MpeV protein; putative [Bacillus subtilis]	59	31	276
131	4	5278	5712	gi 1510655	hypothetical protein (SP:P42297) [Methanococcus jannaschii]	59	39	435
164	1	3	509	gi 1001342	hypothetical protein [Synechocystis sp.]	59	41	507
164	4	1529	2821	gi 1205165	hypothetical protein (SP:P37764) [Haemophilus influenzae]	59	35	1293
164	19	19643	21376	gi 1001381	hypothetical protein [Synechocystis sp.]	59	34	1734
173	3	4727	3717	gi 1184121	auxin-induced protein [Vigna radiata]	59	50	1011
179	2	2218	1688	gi 143036	unidentified gene product [Bacillus subtilis]	59	33	531
195	12	12669	11503	gi 762778	Nifs gene product [Anabaena azollae]	59	41	1167
201	5	4702	5670	gi 1510240	hemin permease [Methanococcus jannaschii]	59	32	969
201	7	5719	6315	gi 1511456	M. jannaschii predicted coding region MJ1437 [Methanococcus jannaschii]	59	34	597
209	1	102	461	gi 1204666	hypothetical protein (GB:X73124_53) [Haemophilus influenzae]	59	42	360
214	3	1050	2234	gi 551531	2-nitropropane dioxygenase [Williopsis saturnus]	59	36	1185
214	5	3293	4135	gi 1303709	YrkJ [Bacillus subtilis]	59	32	843
217	2	3381	2167	gi 290489	dfp (CG Site No. 18430) [Escherichia coli]	59	44	1215
237	5	3078	3785	gi 149382	HisA [Lactococcus lactis]	59	38	708
251	2	376	960	gi 1303791	YqeJ [Bacillus subtilis]	59	34	585
286	1	1621	812	gi 146551	transmembrane protein (kdpD) [Escherichia coli]	59	31	810
316	5	4978	3860	gi 405879	yeiH [Escherichia coli]	59	32	1119
370	3	600	761	gi 1303794	YqeM [Bacillus subtilis]	59	35	162

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
382	1	1009	506	gi1547513	orf3 [Haemophilus influenzae]	59	34	504
391	3	1620	1273	gi1152901	ORF 3 [Spirochaeta aurantia]	59	37	348
406	3	2805	1705	gi1709992	hypothetical protein [Bacillus subtilis]	59	34	1101
426	5	3802	3245	gi11204610	iron(III) dicitrate transport ATP-binding protein FECE [Haemophilus influenzae]	59	36	558
429	2	1513	1148	gi11064809	homologous to sp:HTRA_ECOLI [Bacillus subtilis]	59	42	366
460	2	708	1301	gi1466882	ppsl; B1496_C2_189 [Mycobacterium leprae]	59	37	594
461	4	2212	3135	gi11498295	homoserine kinase homolog [Streptococcus pneumoniae]	59	37	924
473	1	2929	1607	gi1147989	trigger factor [Escherichia coli]	59	40	1323
480	8	5862	6110	gi1205311	(3R)-hydroxymyristol acyl carrier protein dehydrase [Haemophilus influenzae]	59	40	249
521	1	14	1354	pirA25620 A256	staphylocoagulase - Staphylococcus aureus (fragment)	59	32	1341
534	4	2994	4073	gi1153746	mannitol-phosphate dehydrogenase [Streptococcus mutans] pir[C44798 C44798 mannitol-phosphate dehydrogenase Mt1D - treptococcus mutans	59	36	1080
535	1	1	954	gi11469939	group B oligopeptidase PepB [Streptococcus agalactiae]	59	33	954
551	3	2836	3186	gi11204511	bacterioferritin comigratory protein [Haemophilus influenzae]	59	45	351
573	2	449	940	gi1386681	ORF YAL022 [Saccharomyces cerevisiae]	59	36	492
650	1	5	748	gi1396400	similar to eukaryotic Na+/H+ exchangers [Escherichia coli] sp P32703 YJCE_ECOLI HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS NTERGENIC REGION (O549).	59	30	744
664	1	566	285	gi11262748	LukF-PV like component [Staphylococcus aureus]	59	33	282
670	1	3	455	gi11122758	unknown [Bacillus subtilis]	59	42	453
674	3	543	929	gi1293033	integrase [Bacteriophage phi-LC3]	59	46	387
758	1	349	176	gi11500472	M. jannaschii predicted coding region MJ1577 [Methanococcus jannaschii]	59	37	174
771	2	2270	1461	gi1522150	bromoperoxidase BPO-A1 [Streptomyces aureofaciens] sp P33912 BPA1_STRAU NON-HAEM BROMOPEROXIDASE BPO-A1 (EC 1.11.1.-) BROMIDE PEROXIDASE (BPO1). (SUB 2-275)	59	44	810
825	1	2191	1097	gi1397526	clumping factor [Staphylococcus aureus]	59	47	1095
1052	2	1094	723	gi1289262	comE ORF3 [Bacillus subtilis]	59	36	372
1152	1	373	188	gi11276668	ORF238 gene product [Porphyra purpurea]	59	37	186

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1198	1	492	247	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	59	26	246
1441	1	468	235	gi 1045942	glycyl-tRNA synthetase [Mycoplasma genitalium]	59	37	234
2103	1	1	186	gi 459250	triacylglycerol lipase [Galactomyces geotrichum]	59	33	186
2205	1	793	398	gi 1303794	Yqem [Bacillus subtilis]	59	38	396
2578	1	484	284	gi 258003	insulin-like growth factor binding protein complex acid-labile ubunit [rats, liver, Peptide, 603 aa]	59	48	201
2967	2	145	348	gi 1212730	YqhK [Bacillus subtilis]	59	44	204
3012	1	3	248	gi 773571	neurofilament protein NF70 [Helix aspersa]	59	31	246
3544	1	3	401	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	399
3548	1	3	401	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	399
3580	1	698	351	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	348
3720	1	722	363	gi 1408494	homologous to penicillin acylase [Bacillus subtilis]	59	36	360
4171	1	3	296	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	294
4305	1	618	310	gi 1524193	unknown [Mycobacterium tuberculosis]	59	39	309
18	1	1242	622	gi 146913	N-acetylglucosamine transport protein [Escherichia coli] pir B29895 WQEC2N	58	43	621
					phosphotransferase system enzyme II (EC .7.1.69), N-acetylglucosamine-specific - Escherichia coli spiP09323 PTAA ECOLI PTS SYSTEM, N-ACETYLGLUCOSAMINE-SPECIFIC IIABC COMPONENT [EIIA			
20	7	7020	5845	gi 50502	collagen alpha chain precursor (AA -27 to 1127) [Mus musculus]	58	50	1176
21	5	3234	3626	gi 1054860	phosphoribosyl anthranilate isomerase [Thermotoga maritima]	58	32	393
23	2	2841	1669	gi 1276880	EpsG [Streptococcus thermophilus]	58	29	1173
23	10	9301	8090	pir A31133 A311	diaminopimelate decarboxylase (EC 4.1.1.20) - Pseudomonas aeruginosa	58	37	1212
38	29	122555	122884	gi 1973249	vestitone reductase [Medicago sativa]	58	37	330
44	1	2	406	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	58	33	405
45	1	1	552	gi 29464	embryonic myosin heavy chain (1085 AA) [Homo sapiens] ir S12460 S12460	58	33	552
					myosin beta heavy chain - human			
55	2	759	538	gi 158852	glucose regulated protein [Echinococcus multilocularis]	58	32	222
62	13	8493	8068	gi 975353	kinase-associated protein B [Bacillus subtilis]	58	35	426
63	3	1553	1717	gi 166926	[Arabidopsis thaliana unidentified mRNA sequence, complete cds.], ene product [Arabidopsis thaliana]	58	35	165

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
67	13	112017	11229	gi 1228083	NADH dehydrogenase subunit 2 [Chorthippus parallelus]	58	41	789
96	8	8208	9167	gi 709992	hypothetical protein [Bacillus subtilis]	58	42	960
107	2	2065	1364	gi 806327	[Escherichia coli hrpA gene for A protein similar to yeast PRP16 and RP22 [Escherichia coli]	58	37	702
112	7	4519	5613	gi 155588	glucose-fructose oxidoreductase [Zymomonas mobilis] pir A42289 A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis	58	38	1095
114	6	7318	6503	gi 1377843	unknown [Bacillus subtilis]	58	38	816
143	2	2261	1395	pir A45605 A456	mature-parasite-infected erythrocyte surface antigen MESA - Plasmodium falciparum	58	31	867
151	2	717	950	gi 1370261	unknown [Mycobacterium tuberculosis]	58	31	234
154	6	6015	4627	gi 1209277	pCTHom1 gene product [Chlamydia trachomatis]	58	41	1389
154	16	114281	13541	gi 146613	DNA ligase (EC 6.5.1.2) [Escherichia coli]	58	39	741
155	3	2269	1892	gi 1303917	YqiB [Bacillus subtilis]	58	34	378
174	1	1056	529	gi 904198	hypothetical protein [Bacillus subtilis]	58	26	528
189	4	1533	1769	gi 467383	DNA binding protein (probable) [Bacillus subtilis]	58	25	237
201	3	2669	3307	gi 1511453	endonuclease III [Methanococcus jannaschii]	58	34	639
208	1	2	238	gi 1276729	phycobilisome linker polypeptide [Porphyra purpurea]	58	29	237
220	11	114575	13058	gi 397526	clumping factor [Staphylococcus aureus]	58	51	1518
231	3	1629	1474	gi 1002520	MutS [Bacillus subtilis]	58	45	156
233	6	4201	3497	gi 1463023	No definition line found [Caenorhabditis elegans]	58	39	705
243	10	9303	10082	gi 537207	ORF_f277 [Escherichia coli]	58	32	780
257	1	331	1143	gi 1340128	ORF1 [Staphylococcus aureus]	58	44	813
302	2	460	801	gi 40174	ORF X [Bacillus subtilis]	58	34	342
307	11	6984	6127	gi 1303842	YqfU [Bacillus subtilis]	58	30	858
321	3	1914	2747	gi 1239996	hypothetical protein [Bacillus subtilis]	58	41	834
342	4	2724	3497	gi 454838	ORF 6; putative [Pseudomonas aeruginosa]	58	41	774
348	1	1	663	gi 467478	unknown [Bacillus subtilis]	58	36	663
401	2	384	605	gi 143407	para-aminobenzoic acid synthase, component I (pab) [Bacillus ubtilis]	58	53	222

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
437	1	325	1554	gi 1303866	YqgS [Bacillus subtilis]	58	35	1230
445	1	105	1442	gi 581583	protein A [Staphylococcus aureus]	58	32	1338
453	3	789	965	gi 1009455	unknown [Schizosaccharomyces pombe]	58	34	177
453	5	2748	2047	gi 537214	YjjG gene product [Escherichia coli]	58	40	702
479	2	731	1444	gi 1256621	26.7% of identity in 165 aa to a Thermophilic bacterium hypothetical protein 6; putative [Bacillus subtilis]	58	36	714
490	1	909	547	gi 580920	rodD (grtA) polypeptide (AA 1-673) [Bacillus subtilis] pir S06048 S06048	58	36	363
					probable rodD protein - Bacillus subtilis sp P13484 TAGE_BACSU PROBABLE			
					POLY(GLYCEROL-PHOSPHATE) LPHA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS ROTENIN E).			
517	1	1	1164	sp P47264 Y018_	HYPOTHETICAL HELICASE MG018.	58	30	1164
517	6	4182	4544	gi 453422	orf268 gene product [Mycoplasma hominis]	58	29	363
546	3	2802	4019	gi 886052	restriction modification system S subunit [Spiroplasma citri] gi 886052	58	37	1218
					restriction modification system S subunit [Spiroplasma itri]			
562	1	3	179	gi 43831	nifs protein (AA 1-400) [Klebsiella pneumoniae]	58	34	177
600	2	1347	1156	gi 1183839	unknown [Pseudomonas aeruginosa]	58	48	192
604	2	1231	1001	gi 1001353	hypothetical protein [Synechocystis sp.]	58	41	231
619	1	1	504	gi 903748	integral membrane protein [Homo sapiens]	58	43	504
625	1	2	364	gi 1208474	hypothetical protein [Synechocystis sp.]	58	43	363
635	1	1492	755	gi 1510995	transaldolase [Methanococcus jannaschii]	58	41	738
645	1	1	846	gi 677882	ileal sodium-dependent bile acid transporter [Rattus norvegicus] gi 677882	58	33	846
					ileal sodium-dependent bile acid transporter [Rattus orvegicus]			
645	3	906	1556	gi 1239999	hypothetical protein [Bacillus subtilis]	58	41	651
665	1	771	532	gi 1204262	hypothetical protein (GB:ll0328_61) [Haemophilus influenzae]	58	39	240
674	1	635	327	gi 498817	ORF8; homologous to small subunit of phage terminases [Bacillus ubtilis]	58	39	309
675	2	1312	806	gi 42181	osmC gene product [Escherichia coli]	58	28	507
745	1	618	310	gi 1205432	coenzyme PQQ synthesis protein III (pqqIII) [Haemophilus influenzae]	58	32	309
799	2	242	1174	gi 1204669	collagenase [Haemophilus influenzae]	58	36	933
800	2	1096	614	gi 171963	tRNA isopentenyl transferase [Saccharomyces cerevisiae]	58	37	483
					sp P07884 MOD5 YEAST TRNA ISOPENTENYLTRANSFERASE (EC 2.5.1.8) ISOPENTENYL-DIPHOSPHATE: TRNA ISOPENTENYLTRANSFERASE (IPP RANSFERASE) (IPPT).			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
854	1	1108	605	gi 466778	lysine specific permease [Escherichia coli]	58	44	504
885	1	481	242	gi 861199	protoporphyrin IX Mg-chelatase subunit precursor [Hordeum vulgare]	58	33	240
891	1	3	527	gi 1293660	AbsA2 [Streptomyces coelicolor]	58	31	525
942	1	931	467	gi 405567	ttrA [Plasmid pSK41]	58	30	465
1002	1	952	521	gi 577649	preLUM [Staphylococcus aureus]	58	34	432
1438	1	1	261	gi 581558	isoleucyl tRNA synthetase [Staphylococcus aureus] sp P41368 SYIP_STAAU	58	30	261
					ISOLEUCYL-TRNA SYNTHETASE, MUPIROCIN RESISTANT EC 6.1.1.5) (ISOLEUCINE--TRNA LIGASE) (ILERS) (MUPIROCIN RESISTANCE ROTEN).			
1442	1	2	463	gi 971394	similar to Acc.No. D26185 [Escherichia coli]	58	34	462
1873	1	480	241	gi 1339951	small subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	58	38	240
1876	1	3	158	gi 529216	No definition line found [Caenorhabditis elegans] sp P46503 YLY7_CAEEL	58	33	156
					HYPOTHETICAL 7.3 KD PROTEIN F23F12.7 IN HROMOSOME III.			
1989	1	108	401	gi 1405458	YneR [Bacillus subtilis]	58	29	294
2109	1	3	401	gi 1001801	hypothetical protein [Synecocystis sp.]	58	31	399
2473	1	288	145	gi 510140	ligoendopeptidase F [Lactococcus lactis]	58	38	144
2523	1	452	228	gi 644873	catabolic dehydroquinase dehydratase [Acinetobacter calcoaceticus]	58	37	225
3041	1	2	211	gi 1205367	oligopeptide transport ATP-binding protein [Haemophilus influenzae]	58	39	210
3094	1	3	263	gi 1185288	isochorismate synthase [Bacillus subtilis]	58	38	261
3706	1	3	383	gi 456614	mevalonate kinase [Arabidopsis thaliana]	58	48	381
3854	1	1	402	gi 808869	human gcp372 [Homo sapiens]	58	32	402
4082	1	51	224	gi 508551	ribulose-1,5 biphosphate carboxylase large subunit -methyltransferase [Pisum sativum]	58	37	174
4278	1	3	206	gi 180189	cerebellar-degeneration-related antigen (CDR34) [Homo sapiens] gi 182737	58	37	204
					cerebellar degeneration-associated protein [Homo sapiens] pir A29770 A29770 cerebellar degeneration-related protein - human			
19	7	7818	7363	gi 1001516	hypothetical protein [Synecocystis sp.]	57	31	456
23	11	9663	8872	gi 606066	ORF_f256 [Escherichia coli]	57	29	792
31	1	4801	2402	gi 153146	ORF3 [Streptomyces coelicolor]	57	32	2400
38	14	11611	10796	gi 144859	ORF B [Clostridium perfringens]	57	31	816
46	14	112063	13046	gi 1001319	hypothetical protein [Synecocystis sp.]	57	25	984

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
51	3	1411	1187	pir B33856 B338	hypothetical 80K protein - Bacillus sphaericus	57	38	225
54	1	1	453	gil 684950	staphylococcal accessory regulator A [Staphylococcus aureus]	57	31	453
75	1	3	239	gil 1000470	C27B7.7 [Caenorhabditis elegans]	57	42	237
92	5	3855	3061	gil 143607	sporulation protein [Bacillus subtilis]	57	35	795
96	3	4006	4773	gil 144297	acetyl esterase (XynC) [Caldocellum saccharolyticum] pir B37202 B37202	57	34	768
					acetyl esterase (EC 3.1.1.6) (XynC) - Caldocellum accharolyticum			
107	3	1480	2076	gil 460955	TagE [Vibrio cholerae]	57	42	597
109	8	5340	5933	gil 1438846	Unknown [Bacillus subtilis]	57	41	594
112	9	6679	7701	gil 1486250	unknown [Bacillus subtilis]	57	33	1023
114	4	6384	4108	gil 871456	putative alpha subunit of formate dehydrogenase [Methanobacterium hermoautotrophicum]	57	37	2277
126	2	430	1053	gil 288301	ORF2 gene product [Bacillus megaterium]	57	37	624
131	5	6537	6277	gil 1511160	M. jannaschii predicted coding region MJ1163 [Methanococcus jannaschii]	57	38	261
133	3	2668	2201	gil 1303912	Yqhw [Bacillus subtilis]	57	40	468
133	4	3383	2784	gil 1221884	{urea?} amidolyase [Haemophilus influenzae]	57	37	600
147	4	2164	1694	gil 467469	unknown [Bacillus subtilis]	57	33	471
160	2	1293	1060	gil 558604	chitin synthase 2 [Neurospora crassa]	57	28	234
163	8	5687	4764	gil 145580	rarD gene product [Escherichia coli]	57	38	924
168	6	4336	5325	gil 39782	33kDa lipoprotein [Bacillus subtilis]	57	32	990
170	5	3297	3455	gil 603404	Yer164p [Saccharomyces cerevisiae]	57	37	159
221	6	8026	6809	gil 1136221	carboxypeptidase [Sulfolobus solfataricus]	57	32	1218
228	3	1348	1791	gil 288969	fibronectin binding protein [Streptococcus dysgalactiae] pir S33850 S33850	57	32	444
					fibronectin-binding protein - Streptococcus ysgalactiae			
263	4	4411	3686	gil 1185002	dihydrodipicolinate reductase [Pseudomonas syringae pv. tabaci]	57	42	726
276	1	494	255	gil 396380	No definition line found [Escherichia coli]	57	40	240
283	2	335	1324	gil 773349	BirA protein [Bacillus subtilis]	57	32	990
297	1	469	236	gil 1334820	reading frame V [Cauliflower mosaic virus]	57	46	234
342	3	1993	2805	gil 1204431	hypothetical protein (SP:P33644) [Haemophilus influenzae]	57	35	813

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
375	6	3340	3741	gi 385177	cell division protein [Bacillus subtilis]	57	26	402
433	6	3286	4011	gi 1524117	alpha-acetolactate decarboxylase [Lactococcus lactis]	57	40	726
470	3	903	1145	gi 804819	protein serine/threonine kinase [Toxoplasma gondii]	57	30	243
487	5	1391	1723	gi 507323	ORF1 [Bacillus stearothermophilus]	57	28	333
498	1	274	852	gi 1334549	NADH-ubiquinone oxidoreductase subunit 4L [Podospora anserina]	57	34	579
503	1	343	173	gi 1502283	organic cation transporter OCT2 [Rattus norvegicus]	57	30	171
505	2	1619	1284	gi 466884	B1496_C2_194 [Mycobacterium leprae]	57	40	336
519	2	1182	2549	gi 1303707	YrKH [Bacillus subtilis]	57	34	1368
522	2	3234	1945	gi 1064809	homologous to sp:HTRA_ECOLI [Bacillus subtilis]	57	36	1290
538	2	909	1415	gi 153179	phosphorinothyracin n-acetyltransferase [Streptomyces coelicolor] pir JH0246 JH0246 phosphinothricin N-acetyltransferase (EC 2.3.1.-)	57	40	507
					Streptomyces coelicolor			
547	1	968	486	gi 467340	unknown [Bacillus subtilis]	57	50	483
599	1	1062	532	sp P20692 TYRA_	PREPHENATE DEHYDROGENASE (EC 1.3.1.12) (PDH).	57	41	531
620	2	757	572	gi 1107894	unknown [Schizosaccharomyces pombe]	57	38	186
622	2	1600	1130	gi 173028	thioredoxin II [Saccharomyces cerevisiae]	57	39	471
625	2	362	1114	gi 1262366	hypothetical protein [Mycobacterium leprae]	57	34	753
680	1	1	204	gi 143544	RNA polymerase sigma-30 factor [Bacillus subtilis] pir A28625 A28625 transcription initiation factor sigma H - acillus subtilis	57	30	204
690	1	3	629	gi 466520	pocR [Salmonella typhimurium]	57	29	627
696	1	2	433	gi 413972	ipa-48r gene product [Bacillus subtilis]	57	33	432
704	1	36	638	gi 1499931	M. jannaschii predicted coding region MJ1083 [Methanococcus jannaschii]	57	36	603
732	1	2316	1621	gi 1418999	orf4 [Lactobacillus sake]	57	37	696
746	1	451	227	gi 392973	Rab3 [Aplysia californica]	57	42	225
757	1	20	466	gi 43979	L.curvatus small cryptic plasmid gene for rep protein [Lactobacillus rvatus]	57	45	447
862	1	2	295	gi 1303827	Yqfi [Bacillus subtilis]	57	21	294
1049	1	907	455	gi 1510108	ORF-1 [Agrobacterium tumefaciens]	57	35	453
1117	1	1387	695	gi 896286	NH2 terminus uncertain [Leishmania tarentolae]	57	28	693

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1136	1	2	322	gi 1303853	YqgF [Bacillus subtilis]	57	38	321
1144	2	1033	611	gi 310083	voltage-activated calcium channel alpha-1 subunit [Rattus orvegicus]	57	46	423
1172	1	1472	738	gi 1511146	M. jannaschii predicted coding region MJ1143 [Methanococcus jannaschii]	57	28	735
1500	2	746	558	gi 142780	putative membrane protein; putative [Bacillus subtilis]	57	35	189
1676	1	659	399	gi 313777	uracil permease [Escherichia coli]	57	31	261
2481	1	2	400	gi 1237015	ORF4 [Bacillus subtilis]	57	23	399
3099	1	3	230	gi 1204540	isochorismate synthase [Haemophilus influenzae]	57	39	228
3122	1	360	181	gi 882472	ORF_o464 [Escherichia coli]	57	40	180
3560	1	2	361	gi 153490	tetracenomycin C resistance and export protein [Streptomyces laeuscens]	57	37	360
3850	1	856	434	gi 155588	glucose-fructose oxidoreductase [Zymomonas mobilis] pir A42289 A42289	57	40	423
					glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis			
3931	1	704	354	gi 413953	ipa-29d gene product [Bacillus subtilis]	57	36	351
3993	1	1	384	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756	57	39	384
					hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.			
4065	1	793	398	pir JV0037 RDEC	nitrate reductase (EC 1.7.99.4) alpha chain - Escherichia coli	57	31	396
4100	1	596	300	gi 1086633	T06C10.5 gene product [Caenorhabditis elegans]	57	47	297
4163	1	571	287	gi 21512	patatin [Solanum tuberosum]	57	50	285
4267	2	631	335	gi 1000365	SpoIIAG [Bacillus subtilis]	57	38	297
4358	1	3	302	gi 298032	EF [Streptococcus suis]	57	32	300
4389	2	108	290	gi 405894	l-phosphofructokinase [Escherichia coli]	57	37	183
4399	1	2	232	gi 1483603	Pristinamycin I synthase I [Streptomyces pristinaespiralis]	57	35	231
4481	1	572	288	gi 405879	yeiH [Escherichia coli]	57	44	285
4486	1	512	258	gi 515938	glutamate synthase (ferredoxin) [Synechocystis sp.] pir S46957 S46957	57	42	255
					glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.			
4510	1	481	242	gi 1205301	leukotoxin secretion ATP-binding protein [Haemophilus influenzae]	57	38	240
4617	1	468	256	gi 1511222	restriction modification enzyme, subunit M1 [Methanococcus jannaschii]	57	35	213
4	11	12201	11524	gi 149204	histidine utilization repressor G [Klebsiella aerogenes] pir A36730 A36730	56	31	678
					hutG protein - Klebsiella pneumoniae (fragment) sp P19452 HUTG_KLEAE			
					FORMIMINOGLUTAMASE (EC 3.5.3.8) FORMIMINOGLUTAMATE HYDROLASE) (HISTIDINE UTILIZATION PROTEIN G) FRAGMENT)			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
22	8	4248	5177	gi 1322222	RACH1 [Homo sapiens]	56	33	930
38	128	121179	122264	gi 1480705	lipote-protein ligase [Mycoplasma capricolum]	56	34	1086
44	3	1861	2421	gi 490320	Y gene product [unidentified]	56	31	561
44	15	110103	10606	gi 1205099	hypothetical protein (GB:L19201_1) [Haemophilus influenzae]	56	39	504
50	6	4820	5161	gi 209931	fiber protein [Human adenovirus type 5]	56	48	342
53	4	2076	2972	gi 623476	transcriptional activator [Providencia stuartii] sp P43463 AARP_PROST TRANSCRIPTIONAL ACTIVATOR AARP.	56	30	897
67	6	5656	6594	gi 466613	nikB [Escherichia coli]	56	32	939
89	3	2364	1810	gi 482922	protein with homology to pail repressor of B.subtilis [Lactobacillus elbrueckii]	56	39	555
96	1	203	913	gi 145594	cAMP receptor protein (crp) [Escherichia coli]	56	35	711
109	21	18250	17846	gi 1204367	hypothetical protein (GB:U14003_278) [Haemophilus influenzae]	56	27	405
112	8	5611	6678	gi 155588	glucose-fructose oxidoreductase [Zymomonas mobilis] pir A42289 A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis	56	40	1068
131	3	6404	5100	gi 619724	MgtE [Bacillus firmus]	56	30	1305
138	2	65	232	gi 413948	ipa-24d gene product [Bacillus subtilis]	56	31	168
138	4	823	1521	gi 580868	ipa-22r gene product [Bacillus subtilis]	56	31	699
146	2	740	447	gi 1046009	M. genitalium predicted coding region MG309 [Mycoplasma genitalium]	56	37	294
149	2	1639	1067	gi 945380	terminase small subunit [Bacteriophage LL-H]	56	35	573
163	1	2	223	gi 143947	glutamine synthetase [Bacteroides fragilis]	56	30	222
166	5	6745	6449	gi 405792	ORF154 [Pseudomonas putida]	56	26	297
187	1	31	393	gi 311237	H(+)-transporting ATP synthase [Zea mays]	56	30	363
190	1	2	373	gi 1109686	ProX [Bacillus subtilis]	56	35	372
191	8	11538	9943	gi 581070	acyl coenzyme A synthetase [Escherichia coli]	56	35	1596
195	1	1291	647	gi 1510242	collagenase [Methanococcus jannaschii]	56	34	645
230	3	2323	2072	gi 40363	heat shock protein [Clostridium acetobutylicum]	56	39	252
238	5	3383	3775	gi 1477533	sarA [Staphylococcus aureus]	56	31	393
270	2	813	1712	gi 765073	autolysin [Staphylococcus aureus]	56	41	900

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
290	1	3221	1632	gi 547513	orf3 [Haemophilus influenzae]	56	34	1590
297	5	1140	1373	gi 1511556	[M. jannaschii predicted coding region MJ1561 [Methanococcus jannaschii]	56	40	234
321	2	2947	1799	gi 1001801	hypothetical protein [Synechocystis sp.]	56	31	1149
359	2	1279	641	gi 46336	nolI gene product [Rhizobium meliloti]	56	26	639
371	2	360	1823	gi 145304	L-ribulokinase [Escherichia coli]	56	39	1464
391	4	1762	2409	gi 1001634	hypothetical protein [Synechocystis sp.]	56	34	648
402	1	380	192	gi 1438904	5-HT4L receptor [Homo sapiens]	56	48	189
416	4	2480	2109	gi 1408486	HS74A gene product [Bacillus subtilis]	56	31	372
424	3	1756	2334	gi 142471	acetolactate decarboxylase [Bacillus subtilis]	56	32	579
457	1	1907	1017	gi 1205194	[formamidopyrimidine-DNA glycosylase [Haemophilus influenzae]	56	36	891
458	2	2423	1812	gi 15466	terminase [Bacteriophage SPPI]	56	37	612
504	2	2152	1283	gi 1142681	Lpp38 [Pasteurella haemolytica]	56	38	870
511	1	1	1284	gi 217049	brnQ protein [Salmonella typhimurium]	56	37	1284
604	3	1099	1701	gi 467109	rim; 30S Ribosomal protein S18 alanine acetyltransferase; 229_C1_170 [Mycobacterium leprae]	56	43	603
660	5	3547	3774	gi 1229106	ZK930.1 [Caenorhabditis elegans]	56	30	228
707	1	35	400	gi 153929	[NADPH-sulfite reductase flavoprotein component [Salmonella yphimurium]	56	38	366
709	2	1385	1095	gi 1510801	hydrogenase accessory protein [Methanococcus jannaschii]	56	38	291
718	1	1	495	gi 413948	ipa-24d gene product [Bacillus subtilis]	56	35	495
744	1	87	677	gi 928836	repressor protein [Lactococcus lactis phage BK5-T]	56	35	591
790	1	776	399	gi 1511513	ABC transporter, probable ATP-binding subunit [Methanococcus jannaschii]	56	33	378
795	1	3	407	gi 1205382	cell division protein [Haemophilus influenzae]	56	34	405
813	1	19	930	gi 1222161	permease [Haemophilus influenzae]	56	28	912
855	1	3	515	gi 1256621	26.7% of identity in 165 aa to a Thermophilic bacterium hypothetical protein 6; putative [Bacillus subtilis]	56	33	513
968	1	2	466	gi 547513	orf3 [Haemophilus influenzae]	56	37	465
973	2	1049	732	gi 886022	MexR [Pseudomonas aeruginosa]	56	31	318
1203	1	5	223	gi 184251	HMG-1 [Homo sapiens]	56	34	219

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1976	1	452	237	gi 19806	lysine-rich aspartic acid-rich protein [Plasmodium chabaudi] r S22183 S22183 lysine/aspartic acid-rich protein - Plasmodium baudii	56	33	216
2161	1	2	400	gi 1237015	ORF4 [Bacillus subtilis]	56	27	399
2958	1	362	183	gi 466685	No definition line found [Escherichia coli]	56	26	180
2979	1	421	212	gi 1204354	spore germination and vegetative growth protein [Haemophilus influenzae]	56	40	210
2994	2	526	326	gi 836646	phosphoribosylformimino-praic ketoisomerase [Rhodobacter phaeroides]	56	29	201
3026	1	179	328	gi 143306	penicillin V amidase [Bacillus sphaericus]	56	30	150
3189	1	289	146	gi 1166604	Similar to aldehyde dehydrogenase [Caenorhabditis elegans]	56	37	144
3770	1	63	401	gi 1129145	acetyl-CoA C-acyltransferase [Mangifera indica]	56	43	339
4054	2	720	361	gi 1205355	Na+/H+ antiporter [Haemophilus influenzae]	56	31	360
4145	1	1	324	gi 726095	long-chain acyl-CoA dehydrogenase [Mus musculus]	56	36	324
4200	1	505	254	gi 155588	glucose-fructose oxidoreductase [Zymomonas mobilis] pir A42289 A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis	56	40	252
4273	1	675	355	gi 308861	GTG start codon [Lactococcus lactis]	56	33	321
1	3	4095	3436	gi 5341	Putative orf YCLX8c, len:192 [Saccharomyces cerevisiae] r S33591 S33591 hypothetical protein - yeast (Saccharomyces evisiae)	55	25	660
11	12	9377	8505	gi 216773	haloacetate dehalogenase H-1 [Moraxella sp.]	55	32	873
12	4	5133	4534	gi 467337	unknown [Bacillus subtilis]	55	26	600
19	5	5404	5844	gi 1001719	hypothetical protein [Synecocystis sp.]	55	25	441
23	13	14087	12339	gi 474190	iucA gene product [Escherichia coli]	55	30	1749
32	7	5368	6888	gi 1340096	unknown [Mycobacterium tuberculosis]	55	37	1521
34	3	2569	1808	gi 1303968	YqjQ [Bacillus subtilis]	55	39	762
34	5	3960	3412	gi 1303962	YqjK [Bacillus subtilis]	55	33	549
36	1	1291	647	gi 606045	ORF_0118 [Escherichia coli]	55	27	645
36	6	6220	5243	gi 1001341	hypothetical protein [Synecocystis sp.]	55	31	978
47	3	3054	3821	gi 1001819	hypothetical protein [Synecocystis sp.]	55	21	768
49	1	2065	1127	gi 403373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis] pir S37251 S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis	55	36	939

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
67	11	8966	9565	gi 153053	norA199 protein [Staphylococcus aureus]	55	23	600
75	3	881	1273	gi 41698	L-histidinol: NAD+ oxidoreductase (EC 1.1.1.23) (aa 1-434) scherichia coli	55	33	393
82	9	15387	14194	gi 1136221	carboxypeptidase [Sulfolobus solfataricus]	55	35	1194
87	4	3517	4917	gi 1064812	function unknown [Bacillus subtilis]	55	26	1401
88	2	1172	1636	gi 882463	protein-N(pi)-phosphohistidine-sugar phosphotransferase [Escherichia oli]	55	35	465
92	1	127	516	gi 1377832	unknown [Bacillus subtilis]	55	36	390
100	2	836	2035	gi 1370274	zeaxanthin epoxidase [Nicotiana plumbaginifolia]	55	36	1200
100	5	5137	4658	gi 396660	unknown open reading frame [Buchnera aphidicola]	55	29	480
108	3	4266	2986	gi 1499866	M. jannaschii predicted coding region MJ1024 [Methanococcus jannaschii]	55	31	1281
114	3	2616	1834	gi 1511367	formate dehydrogenase, alpha subunit [Methanococcus jannaschii]	55	29	783
144	3	1805	1476	gi 1100787	unkown [Saccharomyces cerevisiae]	55	35	330
165	5	6212	5508	gi 1045884	M. genitalium predicted coding region MG199 [Mycoplasma genitalium]	55	27	705
189	5	2205	2576	gi 142569	ATP synthase a subunit [Bacillus firmus]	55	35	372
191	6	9136	6857	gi 559411	B0272.3 [Caenorhabditis elegans]	55	39	2280
194	2	364	636	gi 1145768	K7 kinesin-like protein [Dictyostelium discoideum]	55	34	273
209	4	1335	1676	gi 473357	thi4 gene product [Schizosaccharomyces pombe]	55	35	342
211	2	1693	1145	gi 410130	ORFX6 [Bacillus subtilis]	55	37	549
213	2	644	1372	gi 633692	TrsA [Yersinia enterocolitica]	55	28	729
214	7	4144	5481	gi 1001793	hypothetical protein [Synecocystis sp.]	55	30	1338
221	7	11473	9197	gi 466520	pocR [Salmonella typhimurium]	55	32	2277
233	8	5908	4817	gi 1237063	unknown [Mycobacterium tuberculosis]	55	38	1092
236	4	1375	2340	gi 1146199	putative [Bacillus subtilis]	55	32	966
243	2	380	1885	gi 459907	mercuric reductase [Plasmid pI258]	55	29	1506
258	1	786	394	gi 455006	orf6 [Rhodococcus fascians]	55	36	393
281	1	126	938	gi 1408493	homologous to SwissProt:YIDA_ECOLI hypothetical protein [Bacillus subtilis]	55	35	813
316	3	1323	2102	gi 1486447	LuxA homologue [Rhizobium sp.]	55	30	780
326	5	2968	2744	gi 1296824	proline iminopeptidase [lactobacillus helveticus]	55	36	225

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
351	2	2322	1429	gi 1204820	hydrogen peroxide-inducible activator [Haemophilus influenzae]	55	28	894
353	4	2197	2412	gi 1272475	chitin synthase [Emericella nidulans]	55	50	216
380	1	14	379	gi 142554	ATP synthase i subunit [Bacillus megaterium]	55	37	366
383	1	462	232	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	55	36	231
386	1	3	938	gi 1510251	DNA helicase, putative [Methanococcus jannaschii]	55	30	936
410	2	1208	1891	gi 1205144	multidrug resistance protein [Haemophilus influenzae]	55	27	684
483	2	411	833	gi 413934	ipa-10r gene product [Bacillus subtilis]	55	26	423
529	3	1777	1433	gi 606150	ORF_f309 [Escherichia coli]	55	33	345
555	1	1088	585	gi 143407	para-aminobenzoic acid synthase, component I (pab) [Bacillus ubtilis]	55	28	504
565	1	402	202	gi 1223961	CDP-tyvelose epimerase [Yersinia pseudotuberculosis]	55	41	201
582	1	751	452	gi 1256643	20.2% identity with NADH dehydrogenase of the Leishmania major mitochondrion; putative [Bacillus subtilis]	55	36	300
645	5	2260	2057	gi 210824	fusion protein F [Bovine respiratory syncytial virus] pir JQ1481 VGNZBA	55	25	204
					fusion glycoprotein precursor - bovine espiratory syncytial virus (strain A51908)			
672	2	957	2216	gi 1511333	M. jannaschii predicted coding region MJ1322 [Methanococcus jannaschii]	55	36	1260
730	1	955	479	gi 537007	ORF_f379 [Escherichia coli]	55	30	477
737	1	1859	945	gi 536963	CG Site No. 18166 [Escherichia coli]	55	30	915
742	2	228	572	gi 304160	product unknown [Bacillus subtilis]	55	38	345
817	2	1211	903	gi 1136289	histidine kinase A [Dictyostelium discoideum]	55	29	309
819	1	582	355	gi 558073	polymorphic antigen [Plasmodium falciparum]	55	22	228
832	2	1152	724	gi 40367	ORFC [Clostridium acetobutylicum]	55	32	429
840	1	769	386	gi 1205875	pseudouridylyate synthase I [Haemophilus influenzae]	55	39	384
1021	1	23	529	gi 48563	beta-lactamase [Yersinia enterocolitica]	55	38	507
1026	1	60	335	gi 47804	Opp C (AA1-301) [Salmonella typhimurium]	55	26	276
1525	1	1	282	gi 1477533	lsrA [Staphylococcus aureus]	55	29	282
1814	2	224	985	gi 1046078	M. genitalium predicted coding region MG369 [Mycoplasma genitalium]	55	38	762
3254	1	427	254	gi 413968	ipa-44d gene product [Bacillus subtilis]	55	30	174

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3695	1	686	345	gi 216773	haloacetate dehalogenase H-1 [Moraxella sp.]	55	32	342
3721	1	1	312	gi 42029	ORF1 gene product [Escherichia coli]	55	31	312
3799	1	3	272	gi 42029	ORF1 gene product [Escherichia coli]	55	38	270
3889	1	22	423	gi 1129145	acetyl-CoA C-acyltransferase [Mangifera indica]	55	45	402
3916	1	2	385	gi 529754	spec [Streptococcus pyogenes]	55	38	384
3945	1	4	198	gi 476252	phase 1 flagellin [Salmonella enterica]	55	36	195
4074	1	488	246	gi 42029	ORF1 gene product [Escherichia coli]	55	38	243
4184	1	2	343	gi 1524267	unknown [Mycobacterium tuberculosis]	55	28	342
4284	1	14	208	gi 1100774	ferredoxin-dependent glutamate synthase [Synechocystis sp.]	55	36	195
4457	2	644	378	gi 180189	cerebellar-degeneration-related antigen (CDR34) [Homo sapiens] gi 182737	55	38	267
					cerebellar degeneration-associated protein [Homo sapiens]			
					pir A29770 A29770 cerebellar degeneration-related protein - human			
4514	1	2	244	gi 216773	haloacetate dehalogenase H-1 [Moraxella sp.]	55	32	243
4599	1	432	217	gi 1129145	acetyl-CoA C-acyltransferase [Mangifera indica]	55	42	216
4606	1	416	210	gi 386120	myosin alpha heavy chain (S2 subfragment) [rabbits, masseter, eptide Partial, 234 aa]	55	27	207
5	8	5348	4932	gi 536069	ORF YBL047c [Saccharomyces cerevisiae]	54	27	417
12	7	7166	6165	gi 1205504	homoserine acetyltransferase [Haemophilus influenzae]	54	30	1002
23	16	17086	15326	gi 474192	lucC gene product [Escherichia coli]	54	31	1761
35	1	2	979	gi 48054	small subunit of soluble hydrogenase (AA 1-384) [Synechococcus sp.]	54	36	978
					ir S06919 HQYC55 soluble hydrogenase (EC 1.12.-.-) small chain -			
					nechococcus sp. (PCC 6716)			
37	11	9437	8667	gi 537207	ORF f277 [Escherichia coli]	54	38	771
37	12	8165	8332	gi 1160967	palmitoyl-protein thioesterase [Homo sapiens]	54	37	168
46	15	13025	13804	gi 438473	protein is hydrophobic, with homology to E. coli ProW; putative Bacillus subtilis]	54	28	780
56	2	203	736	gi 1256139	YbbJ [Bacillus subtilis]	54	34	534
57	13	11117	10179	gi 1151248	inosine-uridine preferring nucleoside hydrolase [Crithidia fasciculata]	54	32	939
66	2	516	1133	gi 1335781	Cap [Drosophila melanogaster]	54	29	618
70	10	8116	8646	gi 1399823	PhoE [Rhizobium meliloti]	54	31	531

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
70	115	12556	11801	sp P02983 TCR_S	TETRACYCLINE RESISTANCE PROTEIN.	54	29	756
87	5	4915	5706	gi 1064811	function unknown [Bacillus subtilis]	54	33	792
92	4	3005	2289	gi 1205366	oligopeptide transport ATP-binding protein [Haemophilus influenzae]	54	33	717
103	2	2596	1556	gi 1710495	protein kinase [Bacillus brevis]	54	33	1041
105	2	3585	2095	gi 1143727	putative [Bacillus subtilis]	54	30	1491
112	4	2337	2732	gi 153724	MalC [Streptococcus pneumoniae]	54	41	396
127	2	1720	2493	gi 1144297	acetyl esterase (XynC) [Caldocellum saccharolyticum] pir B37202 B37202	54	34	774
138	5	1600	3306	gi 142473	pyruvate oxidase [Escherichia coli]	54	36	1707
152	2	525	1172	gi 1377834	unknown [Bacillus subtilis]	54	23	648
161	9	4831	5469	gi 1903305	ORF73 [Bacillus subtilis]	54	28	639
161	113	6694	7251	gi 1511039	phosphate transport system regulatory protein [Methanococcus jannaschii]	54	32	558
164	6	3263	4543	gi 1204976	prolyl-tRNA synthetase [Haemophilus influenzae]	54	34	1281
164	120	121602	22243	gi 1143582	spoIIIEA protein [Bacillus subtilis]	54	32	642
171	6	5683	4250	gi 1436965	[malA] gene products [Bacillus stearothermophilus] pir S43914 S43914	54	37	1434
206	18	19208	19720	gi 1240016	R09E10.3 [Caenorhabditis elegans]	54	38	513
218	2	1090	1905	gi 1467378	unknown [Bacillus subtilis]	54	26	816
220	1	1322	663	gi 11353761	myosin II heavy chain [Naegleria fowleri]	54	22	660
220	113	12655	13059	pir S00485 S004	gene 11-1 protein precursor - Plasmodium falciparum (fragments)	54	35	405
221	3	2030	3709	gi 1303813	YqeW [Bacillus subtilis]	54	34	1680
272	7	5055	4219	gi 162964	arylamine N-acetyltransferase (AA 1-290) [Gallus gallus] ir S06652 XYCHY3	54	33	837
316	7	4141	4701	gi 1682769	lmccE gene product [Escherichia coli]	54	31	561
316	10	6994	8742	gi 1413951	lipa-27d gene product [Bacillus subtilis]	54	28	1749
338	3	3377	2214	gi 1490328	LORF F [unidentified]	54	28	1164
341	4	3201	3614	gi 1171959	myosin-like protein [Saccharomyces cerevisiae]	54	25	414

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
346	1	1820	912	gi 396400	similar to eukaryotic Na+/H+ exchangers [Escherichia coli] sp P32703 YJCE_ECOLI HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS NTERGENIC REGION (O549).	54	34	909
348	2	623	1351	gi 537109	ORF_f343a [Escherichia coli]	54	34	729
378	2	1007	1942	sp P02983 TCR_S	TETRACYCLINE RESISTANCE PROTEIN.	54	31	936
408	6	4351	5301	gi 474190	lucA gene product [Escherichia coli]	54	29	951
444	9	7934	8854	gi 216267	ORF2 [Bacillus megaterium]	54	32	921
463	2	2717	2229	gi 304160	product unknown [Bacillus subtilis]	54	50	489
502	2	1696	1133	gi 1205015	hypothetical protein (SP:P10120) [Haemophilus influenzae]	54	38	564
505	6	6262	5357	gi 1500558	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase [Methanococcus jannaschii]	54	41	906
550	1	2736	1522	gi 40100	rodC (tag3) polypeptide (AA 1-746) [Bacillus subtilis] ir S06049 S06049 rodC protein - Bacillus subtilis p P13485 TAGF_BACSU TECHNOIC ACID BIOSYNTHESIS PROTEIN F.	54	35	1215
551	5	3305	4279	gi 950197	unknown [Corynebacterium glutamicum]	54	34	975
558	2	1356	958	gi 485090	No definition line found [Caenorhabditis elegans]	54	32	399
580	1	91	936	gi 331906	fused envelope glycoprotein precursor [Friend spleen focus-forming irus]	54	45	846
603	3	554	757	gi 1323423	ORF_YGR234w [Saccharomyces cerevisiae]	54	36	204
617	1	25	249	gi 219959	ornithine transcarbamylase [Homo sapiens]	54	40	225
622	3	1097	1480	gi 1303873	YggZ [Bacillus subtilis]	54	25	384
623	1	3	404	gi 1063250	low homology to P20 protein of Bacillus lichiniiformis and bleomycin acetyltransferase of Streptomyces verticillus [Bacillus subtilis]	54	45	402
689	1	1547	1011	gi 552446	NADH dehydrogenase subunit 4 [Apis mellifera ligustica] pir S52968 S52968 NADH dehydrogenase chain 4 - honeybee itochondrion (SGC4)	54	30	537
725	2	686	1441	gi 987096	sensory protein kinase [Streptomyces hygroscopicus]	54	26	756
956	1	1	249	pir S30782 S307	integrin homolog - yeast (Saccharomyces cerevisiae)	54	24	249
978	2	1137	859	gi 1301994	ORF_YNL091w [Saccharomyces cerevisiae]	54	33	279
1314	1	3	281	gi 1001108	hypothetical protein [Synechocystis sp.]	54	33	279
2450	1	1	228	gi 1045057	ch-TOG [Homo sapiens]	54	32	228
2934	1	1	387	gi 580870	ipa-37d qoxA gene product [Bacillus subtilis]	54	36	387
2970	1	499	251	sp P37348 YECE_	HYPOTHETICAL PROTEIN IN ASPS 5'REGION (FRAGMENT).	54	42	249

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3002	1	1	309	gi 44027	Tma protein [Lactococcus lactis]	54	33	309
3561	1	9	464	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	54	35	456
3572	1	72	401	gi 450688	hsdM gene of EcoprrI gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	54	36	330
3829	1	798	400	gi 1322245	mevalonate pyrophosphate decarboxylase [Rattus norvegicus]	54	29	399
3909	1	1	273	gi 29865	CENP-E [Homo sapiens]	54	30	273
3921	1	3	209	pir S24325 S243	glucan 1,4-beta-glucosidase (EC 3.2.1.74) - Pseudomonas fluorescens subsp. cellulosa	54	34	207
4438	1	566	285	gi 1196657	unknown protein [Mycoplasma pneumoniae]	54	30	282
4459	1	3	272	gi 1046081	hypothetical protein (GB:D26185_10) [Mycoplasma genitalium]	54	38	270
4564	1	3	221	gi 216267	ORF2 [Bacillus megaterium]	54	38	219
23	12	12538	10685	gi 474192	iucC gene product [Escherichia coli]	53	35	1854
23	14	14841	13579	gi 42029	ORF1 gene product [Escherichia coli]	53	32	1263
24	3	4440	3940	gi 1369947	c2 gene product [Bacteriophage B1]	53	36	501
26	4	3818	4618	gi 1486247	unknown [Bacillus subtilis]	53	37	801
38	6	2856	3998	gi 405880	yeiI [Escherichia coli]	53	40	1143
38	10	9380	7806	gi 1399954	thyroid sodium/iodide symporter NIS [Rattus norvegicus]	53	29	1575
56	10	12324	12100	pir A54592 A545	110k actin filament-associated protein - chicken	53	32	225
57	6	5047	4583	pir A00341 DEZP	alcohol dehydrogenase (EC 1.1.1.1) - fission yeast (Schizosaccharomyces pombe)	53	39	465
57	12	10515	8932	gi 1480429	putative transcriptional regulator [Bacillus stearothermophilus]	53	30	1584
67	12	9496	10218	gi 1511555	quinolone resistance norA protein protein [Methanococcus jannaschii]	53	31	723
69	3	3125	2382	gi 1087017	arabinogalactan-protein, AGP [Nicotiana glauca, cell-suspension culture filtrate, Peptide, 461 aa]	53	30	744
79	1	3	1031	gi 1523802	glucanase [Anabaena variabilis]	53	32	1029
80	1	673	338	gi 452428	ATPase 3 [Plasmodium falciparum]	53	36	336
88	4	1910	2524	gi 537034	ORF_o488 [Escherichia coli]	53	25	615
88	5	2467	3282	gi 537034	ORF_o488 [Escherichia coli]	53	29	816

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
92	8	5870	5505	gi 399598	amphotropic murine retrovirus receptor [Rattus norvegicus]	53	33	366
94	5	4417	3239	gi 173038	tropomyosin (TPM1) [Saccharomyces cerevisiae]	53	25	1179
99	5	4207	5433	sp P28246 BCR_E	BICYCLOMYCIN RESISTANCE PROTEIN (SULFONAMIDE RESISTANCE PROTEIN).	53	30	1227
120	3	1639	2262	gi 576655	ORF1 [Vibrio anguillarum]	53	35	624
120	11	7257	8897	gi 1524397	glycine betaine transporter OpuD [Bacillus subtilis]	53	33	1641
127	6	6893	5685	gi 1256630	putative [Bacillus subtilis]	53	32	1209
147	2	255	557	gi 581648	epiB gene product [Staphylococcus epidermidis]	53	34	303
158	4	4705	4256	gi 151004	mucooidy regulatory protein AlgR [Pseudomonas aeruginosa] pir A32802 A32802 regulatory protein algR - Pseudomonas aeruginosa sp P26275 ALGR_PSEAE POSITIVE ALGINATE BIOSYNTHESIS REGULATORY ROTEIN.	53	32	450
171	7	5717	5421	gi 1510669	hypothetical protein (GP:D64044_18) [Methanococcus jannaschii]	53	34	297
191	9	13087	11483	gi 298085	acetoacetate decarboxylase [Clostridium acetobutylicum] pir B49346 B49346 butyrate--acetoacetate CoA-transferase (EC .8.3.9) small chain - Clostridium acetobutylicum sp P33752 CTFA_CLOAB BUTYRATE-ACETOACETATE COA-TRANSFERASE SUBUNIT (EC 2.8.3.9) (COAT A)	53	31	1605
203	5	3763	4326	gi 143456	rpoE protein (ttg start codon) [Bacillus subtilis]	53	29	564
206	17	18204	18971	gi 304136	acetylglutamate kinase [Bacillus stearothermophilus] sp Q07905 ARGB_BACST ACETYLGLUTAMATE KINASE (EC 2.7.2.8) (NAG INASE) (AGK) (N-ACETYL-L-GLUTAMATE 5-PHOSPHOTRANSFERASE).	53	36	768
212	10	4021	4221	gi 9878	protein kinase [Plasmodium falciparum]	53	28	201
231	2	1580	1350	gi 537506	paramyosin [Dirofilaria immitis]	53	34	231
272	6	2719	3249	pir A33141 A331	hypothetical protein (gtfd 3' region) - Streptococcus mutans	53	34	531
308	3	927	2576	gi 606292	ORF_o696 [Escherichia coli]	53	33	1650
320	7	5645	5884	gi 160596	RNA polymerase III largest subunit [Plasmodium falciparum] sp P27625 RPC1_PLAFA DNA-DIRECTED RNA POLYMERASE III LARGEST UBUNIT (EC 2.7.7.6).	53	33	240
327	1	218	901	gi 854601	unknown [Schizosaccharomyces pombe]	53	31	684
341	2	212	2500	gi 633732	ORF1 [Campylobacter jejuni]	53	31	2289
351	1	763	383	sp P31675 YABM	HYPOTHETICAL 42.7 KD PROTEIN IN TBPA-LEUD INTERGENIC REGION (ORF104).	53	32	381
433	7	5087	4731	gi 1001961	MHC class II analog [Staphylococcus aureus]	53	30	357
454	2	1240	980	pir A60328 A603	40K cell wall protein precursor (sr 5' region) - Streptococcus mutans (strain OM2175, serotype f)	53	27	261

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
470	4	1123	1761	gi1516826	rat GCP360 [Rattus rattus]	53	30	639
483	1	432	217	gi11480429	putative transcriptional regulator [Bacillus stearothermophilus]	53	33	216
544	1	516	1259	gi146587	ORF 1 (AA 1 - 121) (1 is 2nd base in codon) [Staphylococcus aureus]	53	38	744
				ir S15765 S15765 hypothetical protein 1 (h1b 5' region) - aphylococcus aureus (fragment)				
558	10	3957	3754	gi115140	res gene [Bacteriophage P1]	53	32	204
603	2	339	620	gi1507738	Hmp [Vibrio parahaemolyticus]	53	26	282
693	1	1669	941	gi1153123	toxic shock syndrome toxin-1 precursor [Staphylococcus aureus]	53	38	729
				pir A24606 XCSAS1 toxic shock syndrome toxin-1 precursor - taphylococcus aureus				
766	1	2	673	gi1687600	orfA2; orfA2 forms an operon with orfA1 [Listeria monocytogenes]	53	43	672
781	1	667	335	gi11204551	pilin biogenesis protein [Haemophilus influenzae]	53	26	333
801	1	3	545	gi11279400	SapA protein [Escherichia coli]	53	25	543
803	1	2	910	gi1695278	lipase-like enzyme [Alcaligenes eutrophus]	53	30	909
872	1	1177	590	gi1298032	EF [Streptococcus suis]	53	30	588
910	1	2	184	gi1104936	unknown [Schizosaccharomyces pombe]	53	29	183
943	1	794	399	gi1290508	similar to unidentified ORF near 47 minutes [Escherichia coli]	53	30	396
				sp P31436 YICK_ECOLI HYPOTHETICAL 43.5 KD PROTEIN IN SELC-NLPA NTERGENIC REGION.				
988	1	1004	504	gi1142441	ORF 3; putative [Bacillus subtilis]	53	28	501
1064	1	3	434	gi1305080	myosin heavy chain [Entamoeba histolytica]	53	26	432
1366	1	3	452	gi1308852	transmembrane protein [Lactococcus lactis]	53	33	450
1758	1	792	397	gi11001774	hypothetical protein [Synecocystis sp.]	53	30	396
1897	1	1	447	gi11303949	YqiX [Bacillus subtilis]	53	27	447
2381	1	798	400	gi11146243	22.4% identity with Escherichia coli DNA-damage inducible protein ...; putative [Bacillus subtilis]	53	37	399
3537	1	1	327	gi1450688	hsdM gene of EcooprI gene product [Escherichia coli] pir S38437 S38437 hsdM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	53	35	327
3747	2	137	397	gi11477486	transposase [Burkholderia cepacia]	53	53	261
11	5	3049	3441	gi1868224	No definition line found [Caenorhabditis elegans]	52	33	393

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
15	5	2205	2369	gi 215966	G41 protein (gtg start codon) [Bacteriophage T4]	52	34	165
19	3	2429	3808	gi 1205379	UDP-murac-pentapeptide synthetase [Haemophilus influenzae]	52	31	1380
24	1	6920	3462	gi 579124	predicted 86.4kd protein; 52Kd observed [Mycobacteriophage L5]	52	32	3459
				pir S30971 S30971 gene 26 protein - Mycobacterium phage L5				
				sp Q05233 VG26_BPML5 MINOR TAIL PROTEIN GP26. (SUB 2-837)				
37	5	3015	3935	gi 1500543	PI15 protein [Methanococcus jannaschii]	52	25	921
38	13	8795	9703	gi 46851	glucose kinase [Streptomyces coelicolor]	52	29	909
44	16	110617	11066	gi 42012	moaE gene product [Escherichia coli]	52	36	450
46	1	3	521	gi 1040957	NADH dehydrogenase subunit 6 [Anopheles trinkae]	52	25	519
51	10	5531	6280	gi 388269	traC [Plasmid pAD1]	52	32	750
56	5	3968	2826	gi 181949	endothelial differentiation protein (edg-1) [Homo sapiens]	52	23	1143
				pir A35300 A35300 G protein-coupled receptor edg-1 - human				
				sp P21453 EDG1_HUMAN PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.				
57	5	4850	4173	gi 304153	sorbitol dehydrogenase [Bacillus subtilis]	52	27	678
62	5	3364	2870	gi 1072399	phaE gene product [Rhizobium meliloti]	52	25	495
62	6	4445	3651	gi 46485	NADH dehydrogenase [Synechococcus PCC7942]	52	27	795
67	14	11355	12962	gi 1511365	glutamate synthase (NADPH), subunit alpha [Methanococcus jannaschii]	52	30	1608
67	21	116935	18158	gi 1204393	hypothetical protein (SP:P31122) [Haemophilus influenzae]	52	25	1224
70	4	2185	1997	gi 7227	cytoplasmic dynein heavy chain [Dictyostelium discoideum] r A44357 A44357	52	36	189
				dynein heavy chain, cytosolic - slime mold ctystostellium discoideum)				
96	10	10005	10664	gi 1408485	B65G gene product [Bacillus subtilis]	52	26	660
103	5	3986	3351	gi 1009368	Respiratory nitrate reductase [Bacillus subtilis]	52	42	636
109	3	4102	3350	gi 699274	lmbE gene product [Mycobacterium leprae]	52	39	753
109	19	15732	17300	gi 1526981	amino acid permease YeeF like protein [Salmonella typhimurium]	52	30	1569
121	3	1412	981	gi 732931	unknown [Saccharomyces cerevisiae]	52	32	432
125	3	865	1680	gi 1296975	puT gene product [Porphyromonas gingivalis]	52	38	816
130	2	659	1807	gi 1256634	25.8% identity over 120 aa with the Synenococcus sp. MpeV protein; putative [Bacillus subtilis]	52	36	1149
149	1	1164	583	gi 1225943	PBSX terminase [Bacillus subtilis]	52	33	582
149	14	4687	4415	gi 1510368	M. jannaschii predicted coding region MJ0272 [Methanococcus jannaschii]	52	35	273

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
167	1	216	1001	gi 146025	cell division protein [Escherichia coli]	52	43	786
188	1	120	1256	gi 474915	orf 337; translated orf similarity to SW: BCR_ECOLI bicyclomycin esistance protein of Escherichia coli [Coxiella burnetii] pir S44207 S44207 hypothetical protein 337 - Coxiella burnetii (SUB -338)	52	26	1137
195	9	9161	8760	gi 3028	mitochondrial outer membrane 72K protein [Neurospora crassa] r A36682 A36682 72K mitochondrial outer membrane protein - rospora crassa	52	25	402
200	3	2065	2607	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	52	35	543
203	4	2776	3684	gi 1303698	BltD [Bacillus subtilis]	52	25	909
227	8	5250	5651	gi 305080	myosin heavy chain [Entamoeba histolytica]	52	24	402
242	1	21	1424	gi 1060877	EmrY [Escherichia coli]	52	32	1404
249	5	4526	4753	pir C37222 C372	cytochrome P450 1A1, hepatic - dog (fragment)	52	23	228
255	1	2107	1055	gi 143290	penicillin-binding protein [Bacillus subtilis]	52	28	1053
276	7	3963	3664	gi 1001610	hypothetical protein [Synecocystis sp.]	52	30	300
276	8	4456	4055	gi 416235	orf L3 [Mycoplasma capricolum]	52	26	402
289	2	1856	1449	gi 150900	GTP phosphohydrolase [Proteus vulgaris]	52	34	408
325	1	1	279	gi 1204874	polypeptide deformylase (formylmethionine deformylase) [Haemophilus influenzae]	52	33	279
340	1	2017	1010	gi 1215695	peptide transport system protein SapF homolog; SapF homolog [Mycoplasma pneumoniae]	52	33	1008
375	3	340	1878	gi 467446	similar to SpovB [Bacillus subtilis]	52	28	1539
424	4	4104	3262	gi 1478239	unknown [Mycobacterium tuberculosis]	52	34	843
430	1	3	575	pir A42606 A426	orfA 5' to orf405 - Saccharopolyspora erythraea (fragment)	52	28	573
444	4	4728	3712	gi 1408494	homologous to penicillin acylase [Bacillus subtilis]	52	31	1017
465	1	1802	903	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis] pir A27650 A27650 regulatory protein phoR - Bacillus subtilis sp P23545 PHOR_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOR (EC 2.7.3.-).	52	36	900
469	5	4705	4169	gi 755152	highly hydrophobic integral membrane protein [Bacillus subtilis] sp P42953 TAGG_BACSU TEICHOIC ACID TRANSLOCATION PERMEASE PROTEIN AGG.	52	32	537
495	1	1262	633	gi 1204607	transcription activator [Haemophilus influenzae]	52	25	630
505	7	6004	5762	gi 142440	ATP-dependent nuclease [Bacillus subtilis]	52	28	243

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
517	2	1162	1614	gi 166162	Bacteriophage phi-11 int gene activator [Staphylococcus acteriophage phi 11]	52	35	453
543	2	444	1295	gi 1215693	putative orf; GT9_orf434 [Mycoplasma pneumoniae]	52	25	852
586	1	1	336	gi 581648	epiB gene product [Staphylococcus epidermidis]	52	36	336
773	1	848	426	gi 1279769	FdhC [Methanobacterium thermoformicum]	52	30	423
1120	2	100	330	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	52	35	231
1614	1	691	347	gi 289262	comE ORF3 [Bacillus subtilis]	52	28	345
2495	1	1	324	gi 216151	DNA polymerase (gene L; ttg start codon) [Bacteriophage SPO2] gi 579197 SPO2 DNA polymerase (aa 1-648) [Bacteriophage SPO2] pir A21498 DJBPS2 DNA-directed DNA polymerase (EC 2.7.7.7) - phage P02	52	34	324
2931	1	566	285	gi 1256136	YbbG [Bacillus subtilis]	52	30	282
2943	1	577	320	gi 41713	hisa ORF (AA 1-245) [Escherichia coli]	52	35	258
2993	1	588	295	gi 298032	EF [Streptococcus suis]	52	34	294
3667	1	612	307	gi 849025	hypothetical 64.7-kDa protein [Bacillus subtilis]	52	36	306
3944	1	478	260	gi 1218040	BAA [Bacillus licheniformis]	52	36	219
3954	2	613	347	gi 854064	U87 [Human herpesvirus 6]	52	50	267
3986	1	90	401	gi 1205919	Na+ and Cl- dependent gamma-aminobutyric acid transporter [Haemophilus influenzae]	52	33	312
4002	1	3	389	gi 40003	oxoglutarate dehydrogenase (NADP+) [Bacillus subtilis] p P23129 ODO1_BACSU 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (EC 2.4.2) (ALPHA- KETOGLUTARATE DEHYDROGENASE).	52	42	387
4020	1	1	249	gi 159388	ornithine decarboxylase [Leishmania donovani]	52	47	249
4098	1	438	220	gi 409795	No definition line found [Escherichia coli]	52	32	219
4248	1	3	212	gi 965077	Adr6p [Saccharomyces cerevisiae]	52	40	210
7	1	3	575	gi 895747	putative cel operon regulator [Bacillus subtilis]	51	28	573
21	4	2479	3276	gi 1510962	indole-3-glycerol phosphate synthase [Methanococcus jannaschii]	51	32	798
22	9	5301	5966	gi 1303933	YqjN [Bacillus subtilis]	51	25	666
43	3	1516	1283	gi 1519460	Srp1 [Schizosaccharomyces pombe]	51	31	234
44	17	11042	11305	gi 42011	moaD gene product [Escherichia coli]	51	35	264
51	11	6453	6731	gi 495471	vacuolating toxin [Helicobacter pylori]	51	37	279

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
52	4	2537	2995	gi 1256652	25% identity to the E.coli regulatory protein MprA; putative [Bacillus subtilis]	51	32	459
57	10	7331	6843	gi 508173	EIRA domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	51	32	489
59	1	29	1111	gi 299163	alanine dehydrogenase [Bacillus subtilis]	51	33	1083
67	20	115791	16576	gi 1510977	M. jannaschii predicted coding region MJ0938 [Methanococcus jannaschii]	51	24	786
69	2	1559	1218	gi 467359	unknown [Bacillus subtilis]	51	34	342
71	1	3	1196	gi 298032	EF [Streptococcus suis]	51	32	1194
78	2	349	176	gi 1161242	proliferating cell nuclear antigen [Styela clava]	51	28	174
99	4	3357	4040	gi 642795	TTFID subunit TAFII55 [Homo sapiens]	51	25	684
109	1	2852	1428	gi 580920	rodD (gtAA) polypeptide (AA 1-673) [Bacillus subtilis] pir S06048 S06048	51	27	1425
					probable rodD protein - Bacillus subtilis sp P13484 TAGE BACSU PROBABLE			
					POLY(GLYCEROL-PHOSPHATE) LPHA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS ROtein E).			
109	9	6007	6693	gi 1204815	hypothetical protein (SP:P32662) [Haemophilus influenzae]	51	23	687
112	3	1066	2352	pir S05330 S053	maltose-binding protein precursor - Enterobacter aerogenes	51	42	1287
112	13	14432	12855	gi 405857	yehU [Escherichia coli]	51	29	1578
114	9	9725	8967	gi 435098	orf1 [Mycoplasma capricolum]	51	30	759
115	1	1	912	gi 1431110	ORF YDL085w [Saccharomyces cerevisiae]	51	25	912
127	10	9647	10477	gi 1204314	H. influenzae predicted coding region HI0056 [Haemophilus influenzae]	51	37	831
152	9	6814	7356	gi 431929	MunI regulatory protein [Mycoplasma sp.]	51	38	543
154	2	575	1153	gi 1237044	unknown [Mycobacterium tuberculosis]	51	36	579
154	7	6587	5634	gi 409286	bmrU [Bacillus subtilis]	51	27	954
171	8	6943	6236	gi 1205484	hypothetical protein (SP:P33918) [Haemophilus influenzae]	51	32	708
184	1	1	291	gi 466886	B1496_C3_206 [Mycobacterium leprae]	51	33	291
212	5	1501	2139	pir A45605 A456	mature-parasite-infected erythrocyte surface antigen MESA - Plasmodium falciparum	51	23	639
228	2	707	1378	gi 8204	nuclear protein [Drosophila melanogaster]	51	27	672
236	8	8137	7481	gi 49272	Asparaginase [Bacillus licheniformis]	51	31	657
243	4	4637	3546	gi 1511102	melvalonate kinase [Methanococcus jannaschii]	51	29	1092

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
257	4	3540	3373	gi 1204579	H. influenzae predicted coding region HI0326 [Haemophilus influenzae]	51	22	168
258	3	2397	1609	gi 160299	glutamic acid-rich protein [Plasmodium falciparum] pir A54514 A54514	51	34	789
					glutamic acid-rich protein precursor - Plasmodium alciparum			
265	5	2419	3591	gi 580841	F1 [Bacillus subtilis]	51	32	1173
298	2	518	748	gi 1336162	SCPB [Streptococcus agalactiae]	51	34	231
316	9	5817	7049	gi 413953	ipa-29d gene product [Bacillus subtilis]	51	39	1233
332	2	3775	2057	gi 1209012	mutS [Thermus aquaticus thermophilus]	51	26	1719
364	4	3816	4991	gi 528991	unknown [Bacillus subtilis]	51	32	1176
440	2	448	684	gi 2819	transferase (GAL10) (AA 1 - 687) [Kluyveromyces lactis] r S01407 XUVKG	51	32	237
					UDPglucose 4-epimerase (EC 5.1.3.2) - yeast uyveromyces marxianus var. lactis)			
495	2	1353	1177	gi 297861	protease G [Erwinia chrysanthemi]	51	41	177
495	3	2287	1718	gi 1513317	serine rich protein [Entamoeba histolytica]	51	25	570
506	1	840	421	gi 455320	cII protein [Bacteriophage P4]	51	33	420
600	1	1474	983	gi 587532	orf, len: 201, CAI: 0.16 [Saccharomyces cerevisiae] pir S48818 S48818	51	30	492
					hypothetical protein - yeast (Saccharomyces erevisiae)			
607	3	479	934	gi 1511524	hypothetical protein (SP:P37002) [Methanococcus jannaschii]	51	40	456
686	2	127	600	gi 493017	endocarditis specific antigen [Enterococcus faecalis]	51	30	474
726	1	33	230	gi 1353851	unknown [Prochlorococcus marinus]	51	45	198
861	1	176	652	gi 410145	dehydroquinase dehydratase [Bacillus subtilis]	51	34	477
869	1	782	393	gi 40100	rodC (tag3) polypeptide (AA 1-746) [Bacillus subtilis] ir S06049 S06049	51	23	390
					rodC protein - Bacillus subtilis p P13485 TAGF_BACSU TECHOIC ACID			
					BIOSYNTHESIS PROTEIN F.			
1003	1	642	322	gi 1279707	hypothetical phosphoglycerate mutase [Saccharomyces cerevisiae]	51	39	321
1046	2	866	624	gi 510257	glycosyltransferase [Escherichia coli]	51	29	243
1467	1	702	352	gi 1511175	M. jannaschii predicted coding region MJ1177 [Methanococcus jannaschii]	51	32	351
2558	1	457	230	sp P10582 DPOM_	DNA POLYMERASE (EC 2.7.7.7) (S-1 DNA ORF 3).	51	26	228
3003	1	779	399	gi 809543	CbrC protein [Erwinia chrysanthemi]	51	27	381
3604	1	1	399	pir JC4210 JC42	3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35) - mouse	51	37	399
3732	1	2	316	gi 145906	acyl-CoA synthetase [Escherichia coli]	51	33	315

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3791	1	2	274	gi 1061351	semaphorin III family homolog [Homo sapiens]	51	37	273
3995	1	46	336	gi 216346	surfactin synthetase [Bacillus subtilis]	51	38	291
4193	1	612	307	gi 42749	ribosomal protein L12 (AA 1-179) [Escherichia coli] ir S04776 XSECPL peptide N-acetyltransferase rimL (EC 2.3.1.-) - cherichia coli	51	25	306
4539	1	367	185	gi 1408494	homologous to penicillin acylase [Bacillus subtilis]	51	40	183
4562	1	442	239	gi 1458280	coded for by C. elegans cDNA cm01e7; Similar to hydroxymethylglutaryl-CoA synthase [Caenorhabditis elegans]	51	35	204
1	4	3576	4859	gi 559160	GRAIL score: null; cap site and late promoter motifs present pstream; putative [Autographa californica nuclear polyhedrosis irus]	50	44	1284
11	7	4044	5165	gi 1146207	putative [Bacillus subtilis]	50	35	1122
11	13	10509	9496	gi 1208451	hypothetical protein [Synecocystis sp.]	50	39	1014
19	1	2034	1018	gi 413966	ipa-42d gene product [Bacillus subtilis]	50	29	1017
20	11	8586	8407	gi 1323159	ORF YGR103w [Saccharomyces cerevisiae]	50	28	180
24	5	5408	4824	gi 496280	structural protein [Bacteriophage Tuc2009]	50	29	585
34	4	1926	2759	gi 1303966	YqjO [Bacillus subtilis]	50	36	834
38	30	22865	23440	gi 1072179	Similar to dihydroflavonol-4-reductase (maize, petunia, tomato) [Caenorhabditis elegans]	50	32	576
47	2	1705	2976	gi 153015	FemA protein [Staphylococcus aureus]	50	29	1272
56	13	15290	15841	gi 606096	ORF fl67; end overlaps end of ol00 by 14 bases; start overlaps fl174, ther starts possible [Escherichia coli]	50	30	552
57	1	2135	1077	gi 640922	xylitol dehydrogenase [unidentified hemiascomycete]	50	29	1059
58	2	628	1761	gi 143725	putative [Bacillus subtilis]	50	29	1134
88	6	4393	3884	gi 1072179	Similar to dihydroflavonol-4-reductase (maize, petunia, tomato) [Caenorhabditis elegans]	50	32	510
89	5	3700	3356	gi 1276658	ORF174 gene product [Porphyra purpurea]	50	25	345
141	1	3	239	gi 476024	carbamoyl phosphate synthetase II [Plasmodium falciparum]	50	33	237
151	1	186	626	gi 1403441	unknown [Mycobacterium tuberculosis]	50	35	441
166	7	11065	9623	gi 895747	putative cel operon regulator [Bacillus subtilis]	50	32	1443
201	6	5284	5096	gi 160229	circumsporozoite protein [Plasmodium reichenowi]	50	42	189
206	122	130784	29555	gi 1052754	LmrP integral membrane protein [Lactococcus lactis]	50	24	1230

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
211	4	1523	1927	gi 410131	ORFX7 [Bacillus subtilis]	50	29	405
214	4	2411	3295	sp P37348 YECE_	HYPOTHETICAL PROTEIN IN ASPS 5'REGION (FRAGMENT).	50	37	885
228	7	5068	4406	gi 313580	envelope protein [Human immunodeficiency virus type 1] pir S35835 S35835	50	35	663
					envelope protein - human immunodeficiency virus ype 1 (fragment) (SUB 1-77)			
272	2	3048	1723	gi 1408485	B65G gene product [Bacillus subtilis]	50	22	1326
273	2	1616	984	gi 984186	phosphoglycerate mutase [Saccharomyces cerevisiae]	50	28	633
328	2	2507	1605	gi 148896	lipoprotein [Haemophilus influenzae]	50	26	903
332	4	5469	3802	gi 1526547	DNA polymerase family X [Thermus aquaticus]	50	27	1668
342	5	3473	3931	gi 456562	G-box binding factor [Dictyostelium discoideum]	50	35	459
352	1	1478	741	gi 288301	ORF2 gene product [Bacillus megaterium]	50	29	738
408	7	5299	5523	gi 11665	ORF2136 [Marchantia polymorpha]	50	27	225
420	3	650	1825	gi 757842	UDP-sugar hydrolase [Escherichia coli]	50	30	1176
464	1	1	591	gi 487282	Na+ -ATPase subunit J [Enterococcus hirae]	50	29	591
472	2	1418	864	gi 551875	BglR [Lactococcus lactis]	50	23	555
520	1	23	541	gi 567036	CapE [Staphylococcus aureus]	50	27	519
529	1	6	410	gi 1256652	25% identity to the E.coli regulatory protein MprA; putative [Bacillus subtilis]	50	34	405
534	5	7726	6059	gi 295671	selected as a weak suppressor of a mutant of the subunit AC40 of DNA ependant RNA polymerase I and III [Saccharomyces cerevisiae]	50	18	1668
647	1	2990	1497	gi 405568	TraI protein shares sequence similarity with a family of opoisomerases [Plasmid pSK41]	50	31	1494
664	3	1133	711	gi 410007	leukocidin F component [Staphylococcus aureus, MRSA No. 4, Peptide, 23 aa]	50	32	423
678	1	1	627	gi 298032	EF [Streptococcus suis]	50	29	627
755	3	947	1171	gi 150572	cytochrome c1 precursor (EC 1.10.2.2) [Paracoccus denitrificans] gi 45465	50	37	225
					cytochrome c1 (AA 1-450) [Paracoccus denitrificans] pir C29413 C29413			
					ubiquinol--cytochrome-c reductase (EC 1.10.2.2) ytochrome c1 precursor -			
					Paracoccus denitrificans sp P13627 CY1			
827	1	1363	683	gi 142020	heterocyst differentiation protein (Anabaena sp.)	50	21	681
892	1	3	752	gi 1408485	B65G gene product [Bacillus subtilis]	50	27	750
910	2	438	887	gi 1204727	tyrosine-specific transport protein [Haemophilus influenzae]	50	25	450

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
933	1	524	760	gi 1205451	cell division inhibitor [Haemophilus influenzae]	50	32	237
973	1	424	236	gi 886947	orf3 gene product [Saccharomyces cerevisiae]	50	40	189
1009	1	653	429	gi 153727	M protein [group G streptococcus]	50	28	225
1027	1	511	257	gi 413934	lipa-10r gene product [Bacillus subtilis]	50	25	255
1153	2	556	326	gi 773676	ncca [Alcaligenes xyloxydans]	50	36	231
1222	1	798	400	gi 1408485	B65G gene product [Bacillus subtilis]	50	21	399
1350	1	692	399	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	50	32	294
2945	1	366	184	gi 171704	hexaprenyl pyrophosphate synthetase (COQ1) [Saccharomyces erevisiae]	50	34	183
2968	2	1604	804	gi 397526	clumping factor [Staphylococcus aureus]	50	33	801
2998	2	657	394	gi 495696	F54E7.3 gene product [Caenorhabditis elegans]	50	40	264
3046	2	506	306	pir SL3819 S138	acyl carrier protein - Anabaena variabilis (fragment)	50	32	201
3063	1	547	275	gi 474190	luca gene product [Escherichia coli]	50	29	273
3174	1	3	146	gi 151900	alcohol dehydrogenase [Rhodobacter sphaeroides]	50	31	144
3792	1	625	314	gi 1001423	hypothetical protein [Synechocystis sp.]	50	35	312
3800	1	2	262	gi 144733	NAD-dependent beta-hydroxybutyryl coenzyme A dehydrogenase Clostridium acetobutylicum]	50	28	261
3946	1	373	188	gi 576765	cytochrome b [Myrmecia pilosula]	50	38	186
3984	1	578	291	sp P37348 YECE_HYPOTHETICAL PROTEIN IN ASPS 5'REGION (FRAGMENT).		50	37	288
37	10	8250	7885	gi 1204367	hypothetical protein (GB:U14003_278) [Haemophilus influenzae]	49	30	366
46	16	113802	14848	gi 466860	acd; B1308_F1_34 [Mycobacterium leprae]	49	24	1047
59	5	2267	3601	gi 606304	ORF_o462 [Escherichia coli]	49	27	1335
112	18	17884	18615	gi 559502	ND4 protein (AA 1 - 409) [Caenorhabditis elegans]	49	25	732
138	9	6973	7902	gi 303953	esterase [Acinetobacter calcoaceticus]	49	29	930
217	6	4401	5138	gi 496254	fibronectin/fibrinogen-binding protein [Streptococcus pyogenes]	49	31	738
220	12	11803	12657	gi 397526	clumping factor [Staphylococcus aureus]	49	31	855
228	4	1842	2492	pir S23692 S236	hypothetical protein 9 - Plasmodium falciparum	49	24	651
268	1	5016	2614	gi 143047	ORFB [Bacillus subtilis]	49	26	2403

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
271	2	1164	1373	gi 1001257	hypothetical protein [Synechocystis sp.]	49	38	210
300	3	4340	3180	gi 1510796	hypothetical protein (GP:X91006_2) [Methanococcus jannaschii]	49	26	1161
381	1	2281	1142	gi 396301	[matches PS00041: Bacterial regulatory proteins, araC family ignature [Escherichia coli]	49	29	1140
466	1	3	947	gi 1303863	YggP [Bacillus subtilis]	49	26	945
666	1	379	191	gi 633112	ORF1 [Streptococcus sobrinus]	49	29	189
670	2	403	1014	gi 1122758	unknown [Bacillus subtilis]	49	32	612
709	1	1433	795	gi 143830	xpaC [Bacillus subtilis]	49	29	639
831	1	943	473	gi 401786	phosphomannomutase [Mycoplasma pirum]	49	29	471
1052	1	422	213	gi 1303799	Yqen [Bacillus subtilis]	49	21	210
1800	1	342	172	gi 216300	peptidoglycan synthesis enzyme [Bacillus subtilis] sp P37585 MURG_BACSU MURG PROTEIN UPD-N-ACETYLGLUCOSAMINE--N-ACETYLMURAMYL-PENTAPEPTIDE PYROPHOSPHORYL-UNDECAPRENOL N-ACETYGLUCOSAMINE RANSFERASE).	49	28	171
2430	1	2	376	sp P27434 YFGA_HYPOTHETICAL 36.2 KD PROTEIN IN NDK-GCPE INTERGENIC REGION.	49	26	375	
3096	1	542	273	gi 516360	surfactin synthetase [Bacillus subtilis]	49	25	270
32	4	3771	3100	gi 1217963	hepatocyte nuclear factor 4 gamma (HNF4gamma) [Homo sapiens]	48	36	672
38	1	1	609	gi 1205790	H. influenzae predicted coding region HI1555 [Haemophilus influenzae]	48	28	609
45	6	5021	6427	gi 1524267	unknown [Mycobacterium tuberculosis]	48	20	1407
59	14	16346	31096	gi 1197336	Lmp3 protein [Mycoplasma hominis]	48	28	14751
61	1	3	608	gi 1511555	quinolone resistance norA protein protein [Methanococcus jannaschii]	48	30	606
61	3	3311	3646	gi 1303893	YqhL [Bacillus subtilis]	48	29	336
114	1	98	415	gi 671708	[su(s) homolog; similar to Drosophila melanogaster suppressor of able (su(s)) protein, Swiss-Prot Accession Number P22293 Drosophila virilis]	48	25	318
121	1	1131	610	gi 1314584	unknown [Sphingomonas S88]	48	29	522
136	1	2014	1280	gi 1205968	H. influenzae predicted coding region HI1738 [Haemophilus influenzae]	48	23	735
171	10	8220	9557	gi 1208454	hypothetical protein [Synechocystis sp.]	48	34	1338
175	1	3625	1814	gi 396400	[similar to eukaryotic Na+/H+ exchangers [Escherichia coli] sp P32703 YUCE_ECOLI HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS NTERGENIC REGION (O549).	48	29	1812
194	1	2	385	gi 1510493	M. jannaschii predicted coding region MJ0419 [Methanococcus jannaschii]	48	25	384

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
197	1	901	452	gi 1045714	spermidine/putrescine transport ATP-binding protein [Mycoplasma genitalium]	48	25	450
203	1	1	396	gi 940288	protein localized in the nucleoli of pea nuclei; ORF; putative Pisum sativum	48	29	396
204	1	1363	698	gi 529202	No definition line found [Caenorhabditis elegans]	48	25	666
206	120	134815	12760	gi 511490	gramicidin S synthetase 2 [Bacillus brevis]	48	27	7056
212	1	2	166	gi 295899	nucleolin [Xenopus laevis]	48	34	165
220	110	112652	11426	gi 44073	SecY protein [Lactococcus lactis]	48	23	1227
243	6	6450	5491	gi 1184118	mevalonate kinase [Methanobacterium thermoautotrophicum]	48	30	960
264	4	5434	3308	gi 1015903	ORF YJR151c [Saccharomyces cerevisiae]	48	26	2127
441	1	1532	768	gi 142863	replication initiation protein [Bacillus subtilis] pir B26580 B26580 replication initiation protein - Bacillus ubtilis	48	23	765
444	5	3898	5298	gi 145836	putative [Escherichia coli]	48	24	1401
484	2	388	1110	gi 146551	transmembrane protein (kdpD) [Escherichia coli]	48	18	723
542	3	1425	2000	pir S28969 S289	N-carbamoylsarcosine amidohydrolase (EC 3.5.1.59) - Arthrobacter sp.	48	27	576
566	1	3	1019	gi 153490	tetracenomycin C resistance and export protein [Streptomyces lauecscens]	48	24	1017
611	1	2	730	gi 1103507	unknown [Schizosaccharomyces pombe]	48	38	729
624	1	1255	665	gi 144859	ORF B [Clostridium perfringens]	48	26	591
846	1	1014	508	gi 537506	paramyosin [Dirofilaria immitis]	48	27	507
1020	1	66	950	gi 1499876	magnesium and cobalt transport protein [Methanococcus jannaschii]	48	30	885
1227	1	1	174	gi 493730	lipoxxygenase [Pisum sativum]	48	35	174
1266	1	1	405	gi 882452	ORF f211; alternate name yggA; orf5 of X14436 [Escherichia coli] gi 41425 ORF5 (AA 1-197) [Escherichia coli] (SUB 15-211)	48	24	405
2071	1	707	381	gi 1408486	HS74A gene product [Bacillus subtilis]	48	25	327
2398	1	463	233	gi 1500401	reverse gyrase [Methanococcus jannaschii]	48	40	231
2425	1	476	246	pir H48563 H485	G1 protein - fowlpox virus (strain HP444) (fragment)	48	40	231
2432	1	446	225	gi 1353703	Trio [Homo sapiens]	48	33	222
2453	1	794	399	gi 142850	division initiation protein [Bacillus subtilis]	48	29	396
2998	1	469	236	gi 577569	PepV [Lactobacillus delbrueckii]	48	31	234

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3042	1	14	280	gi 945219	mucin [Homo sapiens]	48	35	267
3686	1	1	405	gi 145836	putative [Escherichia coli]	48	25	405
4027	2	492	301	pir S51177 S511	trans-activator protein - Equine infectious anemia virus	48	32	192
4	2	3641	2232	gi 1303989	YqkI [Bacillus subtilis]	47	24	1410
24	2	599	1084	gi 540083	PC4-1 gene product [Bradyzia hygida]	47	28	486
36	10	7524	6925	gi 1209223	esterase [Acinetobacter lwoffii]	47	26	600
43	2	196	1884	gi 1403455	unknown [Mycobacterium tuberculosis]	47	27	1689
44	22	16118	15108	gi 1511555	quinolone resistance norA protein protein [Methanococcus jannaschii]	47	31	1011
69	7	7141	6710	gi 438466	Possible operon with orfG. Hydrophilic, no homologue in the atabase; putative [Bacillus subtilis]	47	29	432
81	4	5022	4279	gi 466882	pps1; BI496_C2_189 [Mycobacterium leprae]	47	24	744
120	12	9135	8863	gi 927340	D9509.27p; CAI: 0.12 [Saccharomyces cerevisiae]	47	38	273
142	1	2022	1174	gi 486143	ORF YKL094w [Saccharomyces cerevisiae]	47	32	849
168	1	2178	1093	gi 1177254	hypothetical EcsB protein [Bacillus subtilis]	47	29	1086
263	1	1884	943	gi 142822	D-alanine racemase cds [Bacillus subtilis]	47	34	942
279	1	1109	561	gi 516608	2 predicted membrane helices, homology with B. subtilis men Orf3 Rowland et. al. unpublished Accession number M74183), approximately 1 minutes on updated Rudd map; putative [Escherichia coli] sp P37355 YFBB_ECOLI	47	31	549
345	2	2620	1676	gi 1204835	hippuricase [Haemophilus influenzae]	47	28	945
389	2	152	400	gi 456562	G-box binding factor [Dictyostelium discoideum]	47	32	249
391	1	1	831	gi 1420856	myo-inositol transporter [Schizosaccharomyces pombe]	47	19	831
404	3	2072	2773	gi 1255425	C3368.2 gene product [Caenorhabditis elegans]	47	17	702
529	5	2145	3107	gi 1303973	YqjV [Bacillus subtilis]	47	29	963
565	2	2321	1257	gi 142824	processing protease [Bacillus subtilis]	47	28	1065
654	1	962	483	gi 243353	ORF 5' of ECRF3 [herpesvirus saimiri HVS, host-squirrel monkey, eptide, 407 aa]	47	23	480
692	1	115	633	gi 150756	40 kDa protein [Plasmid pJM1]	47	25	519
765	1	1634	819	gi 1256621	26.7% of identity in 165 aa to a Thermophilic bacterium hypothetical protein 6; putative [Bacillus subtilis]	47	28	816

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
825	2	211	1023	gi 397526	clumping factor [Staphylococcus aureus]	47	32	813
914	1	1	615	gi 558073	polymorphic antigen [Plasmodium falciparum]	47	29	615
1076	1	1	753	gi 1147557	Aspartate aminotransferase [Bacillus circulans]	47	33	753
1351	1	793	398	gi 755153	ATP-binding protein [Bacillus subtilis]	47	20	396
4192	1	3	293	gi 145836	putative [Escherichia coli]	47	24	291
5	6	4708	4361	gi 305080	myosin heavy chain [Entamoeba histolytica]	46	30	348
11	4	2777	3058	gi 603639	Yel040p [Saccharomyces cerevisiae]	46	28	282
46	11	110518	110300	gi 1246901	ATP-dependent DNA ligase [Candida albicans]	46	28	219
61	4	3941	7930	gi 298032	EF [Streptococcus suis]	46	35	3990
132	4	5028	4093	gi 1511057	hypothetical protein SP:P45869 [Methanococcus jannaschii]	46	25	936
170	4	4719	3652	pir S51910 S519	G4 protein - Sauroleishmania tarentolae	46	26	1068
191	7	9543	8284	gi 1041334	F54D5.7 [Caenorhabditis elegans]	46	25	1260
253	1	1	396	gi 1204449	dihydrolipoamide acetyltransferase [Haemophilus influenzae]	46	35	396
264	3	437	973	gi 180189	cerebellar-degeneration-related antigen (CDR34) [Homo sapiens] gi 182737	46	29	537
					cerebellar degeneration-associated protein [Homo sapiens]			
					pir A29770 A29770 cerebellar degeneration-related protein - human			
273	1	485	285	gi 607573	envelope glycoprotein C2V3 region [Human immunodeficiency virus type 1]	46	35	201
350	1	3	563	gi 537052	ORF_f286 [Escherichia coli]	46	35	561
384	1	2	862	gi 1221884	(urea?) amidolyase [Haemophilus influenzae]	46	31	861
410	4	1876	2490	gi 1110518	proton antiporter efflux pump [Mycobacterium smegmatis]	46	24	615
432	1	2663	1455	gi 1197634	orf4; putative transporter; Method: conceptual translation supplied by author [Mycobacterium smegmatis]	46	27	1209
458	1	2419	1211	gi 15470	portal protein [Bacteriophage SPP1]	46	30	1209
517	5	2477	4192	gi 1523812	orf5 [Bacteriophage A2]	46	23	1716
540	3	1512	1285	gi 215635	pacA [Bacteriophage P1]	46	30	228
587	2	649	1242	gi 537148	ORF_f181 [Escherichia coli]	46	29	594
1218	1	747	391	gi 1205456	single-stranded-DNA-specific exonuclease [Haemophilus influenzae]	46	30	357

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3685	1	1	402	gi 450688	hsdm gene of Ecoprxi gene product [Escherichia coli] pir S38437 S38437 hsdM	46	33	402
					protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)			
4176	1	673	338	gi 1951460	FIM-C.1 gene product [Xenopus laevis]	46	31	336
37	7	4813	5922	gi 1606064	ORF_f408 [Escherichia coli]	45	24	1110
38	16	11699	12004	gi 1452192	protein tyrosine phosphatase (PTP-BAS, type 2) [Homo sapiens]	45	24	306
87	2	1748	2407	gi 1064813	homologous to sp:PHOR_BACSU [Bacillus subtilis]	45	23	660
103	12	114182	13385	gi 1001307	hypothetical protein [Synechocystis sp.]	45	22	798
112	14	114791	13811	gi 1204389	H. influenzae predicted coding region HI0131 [Haemophilus influenzae]	45	23	981
145	4	4483	3461	gi 220578	open reading frame [Mus musculus]	45	20	1023
170	6	6329	4965	gi 238657	AppC=cytochrome d oxidase, subunit I homolog [Escherichia coli, K12, eptide, 514 aa]	45	27	1365
206	2	5230	4346	gi 1222056	aminotransferase [Haemophilus influenzae]	45	27	885
228	1	60	716	gi 160299	glutamic acid-rich protein [Plasmodium falciparum] pir A54514 A54514 glutamic acid-rich protein precursor - Plasmodium alciparum	45	23	657
288	1	2	1015	gi 1255425	C33G8.2 gene product [Caenorhabditis elegans]	45	23	1014
313	3	4339	3128	gi 581140	NADH dehydrogenase [Escherichia coli]	45	30	1212
332	1	914	459	gi 870966	F47A4.2 [Caenorhabditis elegans]	45	20	456
344	1	3	221	gi 171225	kinesin-related protein [Saccharomyces cerevisiae]	45	26	219
441	2	1501	1073	gi 142863	replication initiation protein [Bacillus subtilis] pir B26580 B26580 replication initiation protein - Bacillus ubtilis	45	27	429
672	1	2	982	gi 1511334	M. jannaschii predicted coding region MJ1323 [Methanococcus jannaschii]	45	22	981
763	3	1345	851	gi 606180	ORF_f310 [Escherichia coli]	45	24	495
886	3	379	846	gi 726426	similar to protein kinases and C. elegans proteins F37C12.8 and 37C12.5 [Caenorhabditis elegans]	45	30	468
948	1	3	473	gi 156400	myosin heavy chain (isozyme unc-54) [Caenorhabditis elegans] pir A93958 MWKW myosin heavy chain B - Caenorhabditis elegans sp P02566 MYSB_CAEEL MYOSIN HEAVY CHAIN B (MHC B).	45	25	471
1158	1	2	376	gi 441155	ransmission-blocking target antigen [Plasmodium falciparum]	45	35	375
2551	1	4	285	gi 1276705	ORF287 gene product [Porphyra purpurea]	45	28	282
3967	1	42	374	gi 976025	HrsA [Escherichia coli]	45	28	333

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
52	7	6931	5846	gi 467378	unknown [Bacillus subtilis]	44	22	1086
138	8	6475	6849	gi 173028	thioredoxin II [Saccharomyces cerevisiae]	44	28	375
221	5	7032	5617	gi 153490	tetracenomycin C resistance and export protein [Streptomyces lauecscens]	44	21	1416
252	2	1331	1122	gi 1204989	hypothetical protein (GB:U00022_9) [Haemophilus influenzae]	44	30	210
263	2	3265	2093	gi 1136221	carboxypeptidase [Sulfolobus solfataricus]	44	26	1173
365	4	4963	3524	gi 1296822	orf1 gene product [Lactobacillus helveticus]	44	31	1440
543	3	1315	1833	gi 1063250	low homology to P20 protein of Bacillus licheniformis and bleomycin acetyltransferase of Streptomyces verticillus [Bacillus subtilis]	44	24	519
544	4	3942	4892	gi 1951460	FIM-C.1 gene product [Xenopus laevis]	44	32	951
792	1	1224	613	gi 205680	high molecular weight neurofilament [Rattus norvegicus]	44	28	612
44	18	11303	11911	gi 1511614	molybdopterin-guanine dinucleotide biosynthesis protein A [Methanococcus jannaschii]	43	27	609
59	8	3665	5128	gi 153490	tetracenomycin C resistance and export protein [Streptomyces lauecscens]	43	21	1464
59	10	5536	7527	gi 153022	lipase [Staphylococcus epidermidis]	43	22	1992
99	1	1346	681	gi 1419051	unknown [Mycobacterium tuberculosis]	43	21	666
310	8	9402	12134	gi 397526	clumping factor [Staphylococcus aureus]	43	21	2733
432	3	2782	2303	pir A60540 A605	sporozoite surface protein 2 - Plasmodium yoelii (fragment)	43	29	480
519	3	2547	3122	sp Q06530 DHSU	SULFIDE DEHYDROGENASE (FLAVOCYTOCHROME C) FLAVOPROTEIN CHAIN PRECURSOR (EC 1.8.2.-) (FC) (FCSO)	43	23	576
4	13	12053	13321	gi 295671	selected as a weak suppressor of a mutant of the subunit AC40 of DNA ependant RNA polymerase I and III [Saccharomyces cerevisiae]	42	18	1269
94	2	1768	1091	gi 501027	ORF2 [Trypanosoma brucei]	42	31	678
127	4	5791	4550	gi 42029	ORF1 gene product [Escherichia coli]	42	21	1242
297	3	1515	1036	gi 142790	ORF1; putative [Bacillus firmus]	42	25	480
344	6	4097	3525	gi 40320	ORF 2 (AA 1-203) [Bacillus thuringiensis]	42	30	573
512	1	2167	1115	gi 405957	yeef [Escherichia coli]	42	23	1053
631	1	2434	1223	gi 580920	rodD (gtAA) polypeptide (AA 1-673) [Bacillus subtilis] pir S06048 S06048	42	24	1212
					probable rodD protein - Bacillus subtilis sp P13484 TAGE_BACSU PROBABLE			
					POLY (GLYCEROL-PHOSPHATE) LPHA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC			
					ACID BIOSYNTHESIS ROTENIN E).			

Table 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
685	3	2359	1739	gi 1303784	YqeD [Bacillus subtilis]	42	19	621
4132	1	787	395	gi 1022910	protein tyrosine phosphatase [Dictyostelium discoideum]	42	25	393
86	2	1375	884	gi 309506	spermidine/spermine N1-acetyltransferase [Mus saxicola] pir S43430 S43430 spermidine/spermine N1-acetyltransferase - spiny ouse (Mus saxicola)	41	30	492
191	12	14797	14075	gi 1124957	orf4 gene product [Methanosarcina barkeri]	41	22	723
212	6	2150	3127	gi 115873	observed 35.2kd protein [Mycobacteriophage 15]	41	26	978
213	3	1263	2000	gi 633692	TrsA [Yersinia enterocolitica]	41	18	738
408	4	2625	3386	gi 1197634	orf4; putative transporter; Method: conceptual translation supplied by author [Mycobacterium smegmatis]	41	24	762
542	1	3	1103	gi 457146	rhoptyr protein [Plasmodium yoelii]	41	21	1101
924	1	2	475	pir JH0148 JH01	nucleolin - rat	41	30	474
1562	1	1	402	gi 552184	asparagine-rich antigen Pfa35-2 [Plasmodium falciparum] pir S27826 S27826 asparagine-rich antigen Pfa35-2 - Plasmodium alciparum (fragment)	40	20	402
2395	1	518	261	pir S42251 S422	hypothetical protein 5 - fowlpox virus	40	18	258
4077	1	3	305	gi 1055055	coded for by C. elegans cDNA yk37g1.5; coded for by C. elegans cDNA yk5c9.5; coded for by C. elegans cDNA yk1a9.5; alternatively spliced form of F52C9.8b [Caenorhabditis elegans]	39	21	303
958	1	1003	503	gi 1255425	C33G8.2 gene product [Caenorhabditis elegans]	37	25	501
59	12	8294	10636	gi 535260	STARP antigen [Plasmodium reichenowi]	36	24	2343
63	5	3550	8079	gi 298032	EF [Streptococcus suis]	36	19	4530
544	3	2507	3601	gi 1015903	ORF YJR151c [Saccharomyces cerevisiae]	35	22	1095
63	4	1949	3574	gi 552195	circumsporozoite protein [Plasmodium falciparum] sp P05691 CSP_PLAFL CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).	32	27	1626

Table 2